PREDICTING OUTCOME WITH TAMOXIFEN IN BREAST CANCER

Related Applications

This application claims benefit of priority from U.S. Provisional Patent Application 60/504,087, filed September 19, 2003, which is hereby incorporated by reference in its entirety as if fully set forth.

Field of the Invention

The invention relates to the identification and use of gene expression profiles, or patterns, with clinical relevance to the treatment of breast cancer using tamoxifen. In particular, the invention provides the identities of genes that are correlated with patient survival and breast cancer recurrence in women treated with tamoxifen. The gene expression profiles, whether embodied in nucleic acid expression, protein expression, or other expression formats, may be used to select subjects afflicted with breast cancer who will likely respond positively to tamoxifen treatment as well as those who will likely be non-responsive and thus candidates for other treatments. The invention also provides the identities of three sets of sequences from three genes with expression patterns that are strongly predictive of responsiveness to tamoxifen.

Background of the Invention

Breast cancer is by far the most common cancer among women. Each year, more than 180,000 and 1 million women in the U.S. and worldwide, respectively, are diagnosed with breast cancer. Breast cancer is the leading cause of death for women between ages 50-55, and is the most common non-preventable malignancy in women in the Western Hemisphere. An estimated 2,167,000 women in the United States are currently living with the disease (National Cancer Institute, Surveillance Epidemiology and End Results (NCI SEER) program, *Cancer Statistics Review (CSR)*, www-seer.ims.nci.nih.gov/Publications/CSR1973 (1998)). Based on cancer rates from 1995 through 1997, a report from the National Cancer Institute (NCI) estimates that about 1 in 8 women in the United States (approximately 12.8 percent) will develop breast cancer during her lifetime (NCI's Surveillance, Epidemiology, and End Results Program (SEER) publication *SEER*

Cancer Statistics Review 1973-1997). Breast cancer is the second most common form of cancer, after skin cancer, among women in the United States. An estimated 250,100 new cases of breast cancer are expected to be diagnosed in the United States in 2001. Of these, 192,200 new cases of more advanced (invasive) breast cancer are expected to occur among women (an increase of 5% over last year), 46,400 new cases of early stage (*in situ*) breast cancer are expected to occur among women (up 9% from last year), and about 1,500 new cases of breast cancer are expected to be diagnosed in men (Cancer Facts & Figures 2001 American Cancer Society). An estimated 40,600 deaths (40,300 women, 400 men) from breast cancer are expected in 2001. Breast cancer ranks second only to lung cancer among causes of cancer deaths in women. Nearly 86% of women who are diagnosed with breast cancer are likely to still be alive five years later, though 24% of them will die of breast cancer after 10 years, and nearly half (47%) will die of breast cancer after 20 years.

Every woman is at risk for breast cancer. Over 70 percent of breast cancers occur in women who have no identifiable risk factors other than age (U.S. General Accounting Office. Breast Cancer, 1971-1991: Prevention, Treatment and Research. GAO/PEMD-92-12; 1991). Only 5 to 10% of breast cancers are linked to a family history of breast cancer (Henderson IC, Breast Cancer. In: Murphy GP, Lawrence WL, Lenhard RE (eds). *Clinical Oncology*. Atlanta, GA: American Cancer Society; 1995:198-219).

Each breast has 15 to 20 sections called lobes. Within each lobe are many smaller lobules. Lobules end in dozens of tiny bulbs that can produce milk. The lobes, lobules, and bulbs are all linked by thin tubes called ducts. These ducts lead to the nipple in the center of a dark area of skin called the areola. Fat surrounds the lobules and ducts. There are no muscles in the breast, but muscles lie under each breast and cover the ribs. Each breast also contains blood vessels and lymph vessels. The lymph vessels carry colorless fluid called lymph, and lead to the lymph nodes. Clusters of lymph nodes are found near the breast in the axilla (under the arm), above the collarbone, and in the chest.

Breast tumors can be either benign or malignant. Benign tumors are not cancerous, they do not spread to other parts of the body, and are not a threat to life. They can usually be removed, and in most cases, do not come back. Malignant tumors are cancerous, and can invade and damage nearby tissues and organs. Malignant tumor cells may metastasize, entering the bloodstream or

lymphatic system. When breast cancer cells metastasize outside the breast, they are often found in the lymph nodes under the arm (axillary lymph nodes). If the cancer has reached these nodes, it means that cancer cells may have spread to other lymph nodes or other organs, such as bones, liver, or lungs.

Major and intensive research has been focused on early detection, treatment and prevention. This has included an emphasis on determining the presence of precancerous or cancerous ductal epithelial cells. These cells are analyzed, for example, for cell morphology, for protein markers, for nucleic acid markers, for chromosomal abnormalities, for biochemical markers, and for other characteristic changes that would signal the presence of cancerous or precancerous cells. This has led to various molecular alterations that have been reported in breast cancer, few of which have been well characterized in human clinical breast specimens. Molecular alterations include presence/absence of estrogen and progesterone steroid receptors, HER-2 expression/amplification (Mark HF, et al. HER-2/neu gene amplification in stages I-IV breast cancer detected by fluorescent in situ hybridization. Genet Med; 1(3):98-103 1999), Ki-67 (an antigen that is present in all stages of the cell cycle except G0 and used as a marker for tumor cell proliferation, and prognostic markers (including oncogenes, tumor suppressor genes, and angiogenesis markers) like p53, p27, Cathepsin D, pS2, multi-drug resistance (MDR) gene, and CD31.

Adjuvant tamoxifen (TAM) is the most effective systemic treatment for estrogen receptor positive (ER+) breast cancer. ER and progesterone receptor (PR) expression have been the major clinicopathological predictor for response to TAM. However, up to 40% of ER+ tumors fail to respond or develop resistance to TAM. Therefore, better predictive biomarkers for TAM response may be able to identify patients who are unlikely to benefit from TAM so that additional or alternative therapies may be sought.

van't Veer et al. (Nature 415:530-536, 2002) describe gene expression profiling of clinical outcome in breast cancer. They identified genes expressed in breast cancer tumors, the expression levels of which correlated either with patients afflicted with distant metastases within 5 years or with patients that remained metastasis-free after at least 5 years.

Ramaswamy et al. (Nature Genetics 33:49-54, 2003) describe the identification of a molecular signature of metastasis in primary solid tumors. The genes of the signature were

identified based on gene expression profiles of 12 metastatic adenocarcinoma nodules of diverse origin (lung, breast, prostate, colorectal, uterus) compared to expression profiles of 64 primary adenocarcinomas representing the same spectrum of tumor types from different individuals. A 128 gene set was identified.

Both of the above described approaches, however, utilize heterogeneous populations of cells found in a tumor sample to obtain information on gene expression patterns. The use of such populations may result in the inclusion or exclusion of multiple genes that are differentially expressed in cancer cells. The gene expression patterns observed by the above described approaches may thus provide little confidence that the differences in gene expression are meaningfully associated with breast cancer recurrence or survival.

Citation of documents herein is not intended as an admission that any is pertinent prior art. All statements as to the date or representation as to the contents of documents is based on the information available to the applicant and does not constitute any admission as to the correctness of the dates or contents of the documents.

Summary of the Invention

The present invention relates to the identification and use of gene expression patterns (or profiles or "signatures") which are clinically relevant to breast cancer. In particular, the identities of genes that are correlated with patient survival and breast cancer recurrence are provided. The gene expression profiles, whether embodied in nucleic acid expression, protein expression, or other expression formats, may be used to predict survival of subjects afflicted with breast cancer and the likelihood of breast cancer recurrence.

The invention thus provides for the identification and use of gene expression patterns (or profiles or "signatures") which correlate with (and thus able to discriminate between) patients with good or poor survival outcomes. In one embodiment, the invention provides patterns that are able to distinguish patients with estrogen receptor positive (ER+) breast tumors into those with that are responsive, or likely to be responsive, to tamoxifen (TAM) treatment and those that are non-responsive, or likely to be non-responsive, to TAM treatment. Responsiveness may be viewed in terms of better survival outcomes over time. These patterns are thus able to distinguish patients with ER+ breast tumors into at least two subtypes.

In a first aspect, the present invention provides a non-subjective means for the identification of patients with ER+ breast cancer as likely to have a good or poor survival outcome following TAM treatment by assaying for the expression patterns disclosed herein. Thus where subjective interpretation may have been previously used to determine the prognosis and/or treatment of breast cancer patients, the present invention provides objective gene expression patterns, which may used alone or in combination with subjective criteria to provide a more accurate assessment of ER+ breast cancer patient outcomes or expected outcomes, including survival and the recurrence of cancer, following treatment with TAM. The expression patterns of the invention thus provide a means to determine ER+ breast cancer prognosis. Furthermore, the expression patterns can also be used as a means to assay small, node negative tumors that are not readily assayed by other means.

The gene expression patterns comprise one or more than one gene capable of discriminating between breast cancer outcomes with significant accuracy. The gene(s) are identified as correlated with ER+ breast cancer outcomes such that the levels of their expression are relevant to a determination of the preferred treatment protocols for a patient. Thus in one embodiment, the invention provides a method to determine the outcome of a subject afflicted with ER+ breast cancer by assaying a cell containing sample from said subject for expression of one or more than one gene disclosed herein as correlated with ER+ breast cancer outcomes following TAM treatment.

Gene expression patterns of the invention are identified as described below. Generally, a large sampling of the gene expression profile of a sample is obtained through quantifying the expression levels of mRNA corresponding to many genes. This profile is then analyzed to identify genes, the expression of which are positively, or negatively, correlated, with ER+ breast cancer outcome with TAM treatment. An expression profile of a subset of human genes may then be identified by the methods of the present invention as correlated with a particular outcome. The use of multiple samples increases the confidence which a gene may be believed to be correlated with a particular survival outcome. Without sufficient confidence, it remains unpredictable whether expression of a particular gene is actually correlated with an outcome and also unpredictable whether expression of a particular gene may be successfully used to identify the outcome for a ER+ breast cancer patient.

A profile of genes that are highly correlated with one outcome relative to another may be used to assay an sample from a subject afflicted with ER+ breast cancer to predict the likely responsiveness (or lack thereof) to TAM in the subject from whom the sample was obtained. Such an assay may be used as part of a method to determine the therapeutic treatment for said subject based upon the breast cancer outcome identified.

The correlated genes may be used singly with significant accuracy or in combination to increase the ability to accurately correlating a molecular expression phenotype with an ER+ breast cancer outcome. This correlation is a way to molecularly provide for the determination of survival outcomes as disclosed herein. Additional uses of the correlated gene(s) are in the classification of cells and tissues; determination of diagnosis and/or prognosis; and determination and/or alteration of therapy.

The ability to discriminate is conferred by the identification of expression of the individual genes as relevant and not by the form of the assay used to determine the actual level of expression. An assay may utilize any identifying feature of an identified individual gene as disclosed herein as long as the assay reflects, quantitatively or qualitatively, expression of the gene in the "transcriptome" (the transcribed fraction of genes in a genome) or the "proteome" (the translated fraction of expressed genes in a genome). Identifying features include, but are not limited to, unique nucleic acid sequences used to encode (DNA), or express (RNA), said gene or epitopes specific to, or activities of, a protein encoded by said gene. All that is required is the identity of the gene(s) necessary to discriminate between ER+ breast cancer outcomes and an appropriate cell containing sample for use in an expression assay.

In another embodiment, the invention provides for the identification of the gene expression patterns by analyzing global, or near global, gene expression from single cells or homogenous cell populations which have been dissected away from, or otherwise isolated or purified from, contaminating cells beyond that possible by a simple biopsy. Because the expression of numerous genes fluctuate between cells from different patients as well as between cells from the same patient sample, multiple data from expression of individual genes and gene expression patterns are used as reference data to generate models which in turn permit the identification of individual gene(s), the expression of which are most highly correlated with particular ER+ breast cancer outcomes.

In additional embodiments, the invention provides physical and methodological means for detecting the expression of gene(s) identified by the models generated by individual expression patterns. These means may be directed to assaying one or more aspects of the DNA template(s) underlying the expression of the gene(s), of the RNA used as an intermediate to express the gene(s), or of the proteinaceous product expressed by the gene(s).

In a further embodiments, the gene(s) identified by a model as capable of discriminating between ER+ breast cancer outcomes may be used to identify the cellular state of an unknown sample of cell(s) from the breast. Preferably, the sample is isolated via non-invasive means. The expression of said gene(s) in said unknown sample may be determined and compared to the expression of said gene(s) in reference data of gene expression patterns correlated with ER+ breast cancer outcomes. Optionally, the comparison to reference samples may be by comparison to the model(s) constructed based on the reference samples.

One advantage provided by the present invention is that contaminating, non-breast cells (such as infiltrating lymphocytes or other immune system cells) are not present to possibly affect the genes identified or the subsequent analysis of gene expression to identify the survival outcomes of patients with breast cancer. Such contamination is present where a biopsy is used to generate gene expression profiles.

In a second aspect, the invention provides a non-subjective means based on the expression of three genes, or combinations thereof, for the identification of patients with ER+ breast cancer as likely to have a good or poor survival outcome following TAM treatment. These three genes are members of the expression patterns disclosed herein which have been found to be strongly predictive of clinical outcome following TAM treatment of ER+ breast cancer.

The present invention thus provides gene sequences identified as differentially expressed in ER+ breast cancer in correlation to TAM responsiveness. The sequences of two of the genes display increased expression in ER+ breast cells that respond to TAM treatment (and thus decreased expression in nonresponsive cases). The sequences of the third gene display decreased expression in ER+ breast cells that respond to TAM treatment (and thus increased expression in nonresponsive cases).

The first set of sequences found to be more highly expressed in TAM responsive, ER+ breast cells are those of interleukin 17 receptor B (IL17RB), which has been mapped to human chromosome 3 at 3p21.1. IL17RB is also referred to as interleukin 17B receptor (IL17BR) and sequences corresponding to it, and thus may be used in the practice of the instant invention, are identified by UniGene Cluster Hs.5470.

The second set of sequences found to be more highly expressed in TAM responsive, ER+ breast cells are those of the calcium channel, voltage-dependent, L type, alpha 1D subunit (CACNA1D), which has been mapped to human chromosome 3 at 3p14.3. Sequences corresponding to CACNA1D, and thus may be used in the practice of the instant invention, are identified by UniGene Cluster Hs.399966.

The set of sequences found to be expressed at lower levels in TAM responsive, ER+ breast cells are those of homeobox B13 (HOXB13), which has been mapped to human chromosome 17 at 17q21.2. Sequences corresponding to HOXB13, and thus may be used in the practice of the instant invention, are identified by UniGene Cluster Hs.66731.

The identified sequences may thus be used in methods of determining the responsiveness of a subject's ER+ breast cancer to TAM treatment via analysis of breast cells in a tissue or cell containing sample from a subject. The present invention provides an non-empirical means for determining TAM responsiveness in ER+ patients. This provides advantages over the use of a "wait and see" approach following treatment with TAM. The expression levels of these sequences may also be used as a means to assay small, node negative tumors that are not readily assessed by conventional means.

The expression levels of the identified sequences may be used alone or in combination with other sequences capable of determining responsiveness to TAM treatment. Preferably, the sequences of the invention are used alone or in combination with each other, such as in the format of a ratio of expression levels that can have improved predictive power over analysis based on expression of sequences corresponding to individual genes.

The present invention provides means for correlating a molecular expression phenotype with a physiological response in a subject with ER+ breast cancer. This correlation provides a way to molecularly diagnose and/or determine treatment for a breast cancer afflicted subject. Additional

uses of the sequences are in the classification of cells and tissues; and determination of diagnosis and/or prognosis. Use of the sequences to identify cells of a sample as responsive, or not, to TAM treatment may be used to determine the choice, or alteration, of therapy used to treat such cells in the subject, as well as the subject itself, from which the sample originated.

An assay of the invention may utilize a means related to the expression level of the sequences disclosed herein as long as the assay reflects, quantitatively or qualitatively, expression of the sequence. Preferably, however, a quantitative assay means is preferred. The ability to determine TAM responsiveness and thus outcome of treatment therewith is provided by the recognition of the relevancy of the level of expression of the identified sequences and not by the form of the assay used to determine the actual level of expression. Identifying features of the sequences include, but are not limited to, unique nucleic acid sequences used to encode (DNA), or express (RNA), the disclosed sequences or epitopes specific to, or activities of, proteins encoded by the sequences. Alternative means include detection of nucleic acid amplification as indicative of increased expression levels (IL17RB and CACNA1D sequences) and nucleic acid inactivation, deletion, or methylation, as indicative of decreased expression levels (HOXB13 sequences). Stated differently, the invention may be practiced by assaying one or more aspect of the DNA template(s) underlying the expression of the disclosed sequence(s), of the RNA used as an intermediate to express the sequence(s), or of the proteinaceous product expressed by the sequence(s). As such, the detection of the amount of, stability of, or degradation (including rate) of, such DNA, RNA and proteinaceous molecules may be used in the practice of the invention.

The practice of the present invention is unaffected by the presence of minor mismatches between the disclosed sequences and those expressed by cells of a subject's sample. A non-limiting example of the existence of such mismatches are seen in cases of sequence polymorphisms between individuals of a species, such as individual human patients within *Homo sapiens*. Knowledge that expression of the disclosed sequences (and sequences that vary due to minor mismatches) is correlated with the presence of non-normal or abnormal breast cells and breast cancer is sufficient for the practice of the invention with an appropriate cell containing sample via an assay for expression.

In one embodiment, the invention provides for the identification of the expression levels of the disclosed sequences by analysis of their expression in a sample containing ER+ breast cells. In one preferred embodiment, the sample contains single cells or homogenous cell populations which have been dissected away from, or otherwise isolated or purified from, contaminating cells beyond that possible by a simple biopsy. Alternatively, undissected cells within a "section" of tissue may be used. Multiple means for such analysis are available, including detection of expression within an assay for global, or near global, gene expression in a sample (e.g. as part of a gene expression profiling analysis such as on a microarray) or by specific detection, such as quantitative PCR (Q-PCR), or real time quantitative PCR.

Preferably, the sample is isolated via non-invasive means. The expression of the disclosed sequence(s) in the sample may be determined and compared to the expression of said sequence(s) in reference data of non-normal breast cells. Alternatively, the expression level may be compared to expression levels in normal cells, preferably from the same sample or subject. In embodiments of the invention utilizing Q-PCR, the expression level may be compared to expression levels of reference genes in the same sample.

When individual breast cells are isolated in the practice of the invention, one benefit is that contaminating, non-breast cells (such as infiltrating lymphocytes or other immune system cells) are not present to possibly affect detection of expression of the disclosed sequence(s). Such contamination is present where a biopsy is used to generate gene expression profiles. However, analysis of differential gene expression and correlation to ER+ breast cancer outcomes with both isolated and non-isolated samples, as described herein, increases the confidence level of the disclosed sequences as capable of having significant predictive power with either type of sample.

While the present invention is described mainly in the context of human breast cancer, it may be practiced in the context of breast cancer of any animal known to be potentially afflicted by breast cancer. Preferred animals for the application of the present invention are mammals, particularly those important to agricultural applications (such as, but not limited to, cattle, sheep, horses, and other "farm animals"), animal models of breast cancer, and animals for human companionship (such as, but not limited to, dogs and cats).

Brief Description of the Drawings

Figure 1 shows the survival curves for two groups of breast cancer patients defined by expression signatures based on 149 genes as described herein.

Figure 2 shows survival curves for two groups of breast cancer patients defined by expression signatures based on genes sets identified for whole tissue sections (left graph) and laser microdissected cells (right graph) as described herein.

Figure 3 shows the expression levels of IL17BR, HOXB13, and CACNA1D in whole tissue sections (top three graphs) and laser microdissected cells (bottom three graphs).

Figure 4 shows receiver operating characteristic (ROC) analyses of IL17BR, HOXB13, and CACNA1D expression levels as predictors of breast cancer outcomes in whole tissue sections (top three graphs) and laser microdissected cells (bottom three graphs). AUC refers to area under the curve.

Figure 5 shows Kaplan-Meier (KM) analyses of IL17BR, HOXB13, and CACNA1D expression levels as predictors of breast cancer outcomes in whole tissue sections (top three graphs) and laser microdissected cells (bottom three graphs).

Figure 6 shows expression levels (top three graphs) and ROC (bottom three graphs) analysis of IL17BR, HOXB13, and CACNA1D as predictors of breast cancer outcomes in macrodissected formalin fixed, paraffin embedded (FFPE) samples from a cohort of 31 patients treated with tamoxifen.

Figure 7 shows analysis and use of a ratio of HOXB13 to IL17BR expression levels as a predictor of breast cancer outcome. Plots of the ratios in whole tissue sections and macrodissected FFPE samples as well as ROC analysis are shown in the first four graphs. Survival curves based on "high" and "low" ratios (relative to 0.22, the horizontal line in the plots of the ratios) are shown in the last graph.

Modes of Practicing the Invention

Definitions of terms as used herein:

A gene expression "pattern" or "profile" or "signature" refers to the relative expression of genes correlated with responsiveness to TAM treatment of ER+ breast cancer. Responsiveness or

lack thereof may be expressed as survival outcomes which are correlated with an expression "pattern" or "profile" or "signature" that is able to distinguish between, and predict, said outcomes.

A "gene" is a polynucleotide that encodes a discrete product, whether RNA or proteinaceous in nature. It is appreciated that more than one polynucleotide may be capable of encoding a discrete product. The term includes alleles and polymorphisms of a gene that encodes the same product, or a functionally associated (including gain, loss, or modulation of function) analog thereof, based upon chromosomal location and ability to recombine during normal mitosis.

A "sequence" or "gene sequence" as used herein is a nucleic acid molecule or polynucleotide composed of a discrete order of nucleotide bases. The term includes the ordering of bases that encodes a discrete product (i.e. "coding region"), whether RNA or proteinaceous in nature, as well as the ordered bases that precede or follow a "coding region". Non-limiting examples of the latter include 5' and 3' untranslated regions of a gene. It is appreciated that more than one polynucleotide may be capable of encoding a discrete product. It is also appreciated that alleles and polymorphisms of the disclosed sequences may exist and may be used in the practice of the invention to identify the expression level(s) of the disclosed sequences or the allele or polymorphism. Identification of an allele or polymorphism depends in part upon chromosomal location and ability to recombine during mitosis.

The terms "correlate" or "correlation" or equivalents thereof refer to an association between expression of one or more genes and a physiological response of a breast cancer cell and/or a breast cancer patient in comparison to the lack of the response. A gene may be expressed at higher or lower levels and still be correlated with responsiveness or breast cancer survival or outcome. The invention provides for the correlation between increases in expression of IL17RB and CACNA1D sequences and TAM responsiveness in ER+ breast cells. Similarly, the invention provides for the correlation between decreases in expression of HOXB13 sequences and TAM responsiveness in ER+ breast cells. Increases and decreases may be readily expressed in the form of a ratio between expression in a non-normal cell and a normal cell such that a ratio of one (1) indicates no difference while ratios of two (2) and one-half indicate twice as much, and half as much, expression in the non-normal cell versus the normal cell, respectively. Expression levels can be readily determined by quantitative methods as described below.

For example, increases in IL17RB expression can be indicated by ratios of or about 1.1, of or about 1.2, of or about 1.3, of or about 1.4, of or about 1.5, of or about 1.6, of or about 1.7, of or about 1.8, of or about 1.9, of or about 2, of or about 2.5, of or about 3, of or about 3.5, of or about 4, of or about 4.5, of or about 5, of or about 5.5, of or about 6, of or about 6.5, of or about 7, of or about 7.5, of or about 8, of or about 8.5, of or about 9, of or about 9.5, of or about 10, of or about 15, of or about 20, of or about 30, of or about 40, of or about 50, of or about 60, of or about 70, of or about 80, of or about 90, of or about 100, of or about 150, of or about 200, of or about 300, of or about 400, of or about 500, of or about 600, of or about 700, of or about 800, of or about 900, or of or about 1000. A ratio of 2 is a 100% (or a two-fold) increase in expression. Similar ratios can be used with respect to increases in CACNA1D expression. Decreases in HOXB13 expression can be indicated by ratios of or about 0.9, of or about 0.8, of or about 0.7, of or about 0.6, of or about 0.5, of or about 0.4, of or about 0.3, of or about 0.2, of or about 0.1, of or about 0.05, of or about 0.01, of or about 0.005, of or about 0.001, of or about 0.0005, of or about 0.0001, of or about 0.00005, of or about 0.00001, of or about 0.000005, or of or about 0.000001.

A "polynucleotide" is a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. This term refers only to the primary structure of the molecule. Thus, this term includes double- and single-stranded DNA and RNA. It also includes known types of modifications including labels known in the art, methylation, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, and internucleotide modifications such as uncharged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), as well as unmodified forms of the polynucleotide.

The term "amplify" is used in the broad sense to mean creating an amplification product can be made enzymatically with DNA or RNA polymerases. "Amplification," as used herein, generally refers to the process of producing multiple copies of a desired sequence, particularly those of a sample. "Multiple copies" mean at least 2 copies. A "copy" does not necessarily mean perfect sequence complementarity or identity to the template sequence. Methods for amplifying mRNA are generally known in the art, and include reverse transcription PCR (RT-PCR) and those described in U.S. Patent Application 10/062,857 (filed on October 25, 2001), as well as U.S. Provisional Patent Applications 60/298,847 (filed June 15, 2001) and 60/257,801 (filed December 22, 2000), all of

which are hereby incorporated by reference in their entireties as if fully set forth. Another method which may be used is quantitative PCR (or Q-PCR). Alternatively, RNA may be directly labeled as the corresponding cDNA by methods known in the art.

By "corresponding", it is meant that a nucleic acid molecule shares a substantial amount of sequence identity with another nucleic acid molecule. Substantial amount means at least 95%, usually at least 98% and more usually at least 99%, and sequence identity is determined using the BLAST algorithm, as described in Altschul et al. (1990), J. Mol. Biol. 215:403-410 (using the published default setting, i.e. parameters w=4, t=17).

A "microarray" is a linear or two-dimensional array of preferably discrete regions, each having a defined area, formed on the surface of a solid support such as, but not limited to, glass, plastic, or synthetic membrane. The density of the discrete regions on a microarray is determined by the total numbers of immobilized polynucleotides to be detected on the surface of a single solid phase support, preferably at least about 50/cm^2, more preferably at least about 100/cm^2, even more preferably at least about 500/cm^2, but preferably below about 1,000/cm^2. Preferably, the arrays contain less than about 500, about 1000, about 1500, about 2000, about 2500, or about 3000 immobilized polynucleotides in total. As used herein, a DNA microarray is an array of oligonucleotides or polynucleotides placed on a chip or other surfaces used to hybridize to amplified or cloned polynucleotides from a sample. Since the position of each particular group of primers in the array is known, the identities of a sample polynucleotides can be determined based on their binding to a particular position in the microarray.

Because the invention relies upon the identification of genes that are over- or under-expressed, one embodiment of the invention involves determining expression by hybridization of mRNA, or an amplified or cloned version thereof, of a sample cell to a polynucleotide that is unique to a particular gene sequence. Preferred polynucleotides of this type contain at least about 20, at least about 22, at least about 24, at least about 26, at least about 28, at least about 30, or at least about 32 consecutive basepairs of a gene sequence that is not found in other gene sequences. The term "about" as used in the previous sentence refers to an increase or decrease of 1 from the stated numerical value. Even more preferred are polynucleotides of at least or about 50, at least or about 100, at least about or 150, at least or about 200, at least or about 250, at least or about 300, at least

or about 350, at least or about 400, , at least or about 450, or at least or about 500 consecutive bases of a sequence that is not found in other gene sequences. The term "about" as used in the preceding sentence refers to an increase or decrease of 10% from the stated numerical value. Longer polynucleotides may of course contain minor mismatches (e.g. via the presence of mutations) which do not affect hybridization to the nucleic acids of a sample. Such polynucleotides may also be referred to as polynucleotide probes that are capable of hybridizing to sequences of the genes, or unique portions thereof, described herein. Such polynucleotides may be labeled to assist in their detection. Preferably, the sequences are those of mRNA encoded by the genes, the corresponding cDNA to such mRNAs, and/or amplified versions of such sequences. In preferred embodiments of the invention, the polynucleotide probes are immobilized on an array, other solid support devices, or in individual spots that localize the probes.

In another embodiment of the invention, all or part of a disclosed sequence may be amplified and detected by methods such as the polymerase chain reaction (PCR) and variations thereof, such as, but not limited to, quantitative PCR (Q-PCR), reverse transcription PCR (RT-PCR), and real-time PCR, optionally real-time RT-PCR. Such methods would utilize one or two primers that are complementary to portions of a disclosed sequence, where the primers are used to prime nucleic acid synthesis. The newly synthesized nucleic acids are optionally labeled and may be detected directly or by hybridization to a polynucleotide of the invention. The newly synthesized nucleic acids may be contacted with polynucleotides (containing sequences) of the invention under conditions which allow for their hybridization.

Alternatively, and in yet another embodiment of the invention, gene expression may be determined by analysis of expressed protein in a cell sample of interest by use of one or more antibodies specific for one or more epitopes of individual gene products (proteins) in said cell sample. Such antibodies are preferably labeled to permit their easy detection after binding to the gene product.

The term "label" refers to a composition capable of producing a detectable signal indicative of the presence of the labeled molecule. Suitable labels include radioisotopes, nucleotide chromophores, enzymes, substrates, fluorescent molecules, chemiluminescent moieties, magnetic

particles, bioluminescent moieties, and the like. As such, a label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means.

The term "support" refers to conventional supports such as beads, particles, dipsticks, fibers, filters, membranes and silane or silicate supports such as glass slides.

As used herein, a "breast tissue sample" or "breast cell sample" refers to a sample of breast tissue or fluid isolated from an individual suspected of being afflicted with, or at risk of developing, breast cancer. Such samples are primary isolates (in contrast to cultured cells) and may be collected by any non-invasive means, including, but not limited to, ductal lavage, fine needle aspiration, needle biopsy, the devices and methods described in U.S. Patent 6,328,709, or any other suitable means recognized in the art. Alternatively, the "sample" may be collected by an invasive method, including, but not limited to, surgical biopsy.

"Expression" and "gene expression" include transcription and/or translation of nucleic acid material.

As used herein, the term "comprising" and its cognates are used in their inclusive sense; that is, equivalent to the term "including" and its corresponding cognates.

Conditions that "allow" an event to occur or conditions that are "suitable" for an event to occur, such as hybridization, strand extension, and the like, or "suitable" conditions are conditions that do not prevent such events from occurring. Thus, these conditions permit, enhance, facilitate, and/or are conducive to the event. Such conditions, known in the art and described herein, depend upon, for example, the nature of the nucleotide sequence, temperature, and buffer conditions. These conditions also depend on what event is desired, such as hybridization, cleavage, strand extension or transcription.

Sequence "mutation," as used herein, refers to any sequence alteration in the sequence of a gene disclosed herein interest in comparison to a reference sequence. A sequence mutation includes single nucleotide changes, or alterations of more than one nucleotide in a sequence, due to mechanisms such as substitution, deletion or insertion. Single nucleotide polymorphism (SNP) is also a sequence mutation as used herein. Because the present invention is based on the relative level of gene expression, mutations in non-coding regions of genes as disclosed herein may also be assayed in the practice of the invention.

"Detection" includes any means of detecting, including direct and indirect detection of gene expression and changes therein. For example, "detectably less" products may be observed directly or indirectly, and the term indicates any reduction (including the absence of detectable signal). Similarly, "detectably more" product means any increase, whether observed directly or indirectly.

Increases and decreases in expression of the disclosed sequences are defined in the following terms based upon percent or fold changes over expression in normal cells. Increases may be of 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, or 200% relative to expression levels in normal cells. Alternatively, fold increases may be of 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, or 10 fold over expression levels in normal cells. Decreases may be of 10, 20, 30, 40, 50, 55, 60, 65, 70, 75, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 99 or 100% relative to expression levels in normal cells.

Unless defined otherwise all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this invention belongs.

Embodiments of the Invention

In a first aspect, the disclosed invention relates to the identification and use of gene expression patterns (or profiles or "signatures") which discriminate between (or are correlated with) breast cancer survival in a subject treated with tamoxifen (TAM). Such patterns may be determined by the methods of the invention by use of a number of reference cell or tissue samples, such as those reviewed by a pathologist of ordinary skill in the pathology of breast cancer, which reflect breast cancer cells as opposed to normal or other non-cancerous cells. The outcomes experienced by the subjects from whom the samples may be correlated with expression data to identify patterns that correlate with the outcomes following TAM treatment. Because the overall gene expression profile differs from person to person, cancer to cancer, and cancer cell to cancer cell, correlations between certain cells and genes expressed or underexpressed may be made as disclosed herein to identify genes that are capable of discriminating between breast cancer outcomes.

The present invention may be practiced with any number of the genes believed, or likely to be, differentially expressed with respect to breast cancer outcomes, particularly in cases of ER+ breast cancer. The identification may be made by using expression profiles of various homogenous breast cancer cell populations, which were isolated by microdissection, such as, but not limited to,

laser capture microdissection (LCM) of 100-1000 cells. The expression level of each gene of the expression profile may be correlated with a particular outcome. Alternatively, the expression levels of multiple genes may be clustered to identify correlations with particular outcomes.

Genes with significant correlations to breast cancer survival when the subject is treated with tamoxifen may be used to generate models of gene expressions that would maximally discriminate between outcomes where a subject responds to tamoxifen treatment and outcomes where the tamoxifen treatment is not successful. Alternatively, genes with significant correlations may be used in combination with genes with lower correlations without significant loss of ability to discriminate between outcomes. Such models may be generated by any appropriate means recognized in the art, including, but not limited to, cluster analysis, supported vector machines, neural networks or other algorithm known in the art. The models are capable of predicting the classification of a unknown sample based upon the expression of the genes used for discrimination in the models. "Leave one out" cross-validation may be used to test the performance of various models and to help identify weights (genes) that are uninformative or detrimental to the predictive ability of the models. Cross-validation may also be used to identify genes that enhance the predictive ability of the models.

The gene(s) identified as correlated with particular breast cancer outcomes relating to tamoxifen treatment by the above models provide the ability to focus gene expression analysis to only those genes that contribute to the ability to identify a subject as likely to have a particular outcome relative to another. The expression of other genes in a breast cancer cell would be relatively unable to provide information concerning, and thus assist in the discrimination of, a breast cancer outcome.

As will be appreciated by those skilled in the art, the models are highly useful with even a small set of reference gene expression data and can become increasingly accurate with the inclusion of more reference data although the incremental increase in accuracy will likely diminish with each additional datum. The preparation of additional reference gene expression data using genes identified and disclosed herein for discriminating between different tamoxifen treatment outcomes in breast cancer is routine and may be readily performed by the skilled artisan to permit the

generation of models as described above to predict the status of an unknown sample based upon the expression levels of those genes.

To determine the (increased or decreased) expression levels of genes in the practice of the present invention, any method known in the art may be utilized. In one preferred embodiment of the invention, expression based on detection of RNA which hybridizes to the genes identified and disclosed herein is used. This is readily performed by any RNA detection or amplification+detection method known or recognized as equivalent in the art such as, but not limited to, reverse transcription-PCR, the methods disclosed in U.S. Patent Application 10/062,857 (filed on October 25, 2001) as well as U.S. Provisional Patent Applications 60/298,847 (filed June 15, 2001) and 60/257,801 (filed December 22, 2000), and methods to detect the presence, or absence, of RNA stabilizing or destabilizing sequences.

Alternatively, expression based on detection of DNA status may be used. Detection of the DNA of an identified gene as methylated or deleted may be used for genes that have decreased expression in correlation with a particular breast cancer outcome. This may be readily performed by PCR based methods known in the art, including, but not limited to, Q-PCR. Conversely, detection of the DNA of an identified gene as amplified may be used for genes that have increased expression in correlation with a particular breast cancer outcome. This may be readily performed by PCR based, fluorescent *in situ* hybridization (FISH) and chromosome *in situ* hybridization (CISH) methods known in the art.

Expression based on detection of a presence, increase, or decrease in protein levels or activity may also be used. Detection may be performed by any immunohistochemistry (IHC) based, blood based (especially for secreted proteins), antibody (including autoantibodies against the protein) based, exfoliate cell (from the cancer) based, mass spectroscopy based, and image (including used of labeled ligand) based method known in the art and recognized as appropriate for the detection of the protein. Antibody and image based methods are additionally useful for the localization of tumors after determination of cancer by use of cells obtained by a non-invasive procedure (such as ductal lavage or fine needle aspiration), where the source of the cancerous cells is not known. A labeled antibody or ligand may be used to localize the carcinoma(s) within a patient.

A preferred embodiment using a nucleic acid based assay to determine expression is by immobilization of one or more sequences of the genes identified herein on a solid support, including, but not limited to, a solid substrate as an array or to beads or bead based technology as known in the art. Alternatively, solution based expression assays known in the art may also be used. The immobilized gene(s) may be in the form of polynucleotides that are unique or otherwise specific to the gene(s) such that the polynucleotide would be capable of hybridizing to a DNA or RNA corresponding to the gene(s). These polynucleotides may be the full length of the gene(s) or be short sequences of the genes (up to one nucleotide shorter than the full length sequence known in the art by deletion from the 5' or 3' end of the sequence) that are optionally minimally interrupted (such as by mismatches or inserted non-complementary basepairs) such that hybridization with a DNA or RNA corresponding to the gene(s) is not affected. Preferably, the polynucleotides used are from the 3' end of the gene, such as within about 350, about 300, about 250, about 200, about 150, about 100, or about 50 nucleotides from the polyadenylation signal or polyadenylation site of a gene or expressed sequence. Polynucleotides containing mutations relative to the sequences of the disclosed genes may also be used so long as the presence of the mutations still allows hybridization to produce a detectable signal.

The immobilized gene(s) may be used to determine the state of nucleic acid samples prepared from sample breast cell(s) for which the outcome of the sample's subject (e.g. patient from whom the sample is obtained) is not known or for confirmation of an outcome that is already assigned to the sample's subject. Without limiting the invention, such a cell may be from a patient with ER+ breast cancer. The immobilized polynucleotide(s) need only be sufficient to specifically hybridize to the corresponding nucleic acid molecules derived from the sample under suitable conditions. While even a single correlated gene sequence may to able to provide adequate accuracy in discriminating between two breast cancer outcomes, two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, nine or more, ten or more, or eleven or more of the genes identified herein may be used as a subset capable of discriminating may be used in combination to increase the accuracy of the method. The invention specifically contemplates the selection of more than one, two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, nine or more, ten or more, or eleven or more of the genes disclosed in

the tables and figures herein for use as a subset in the identification of breast cancer survival outcome.

Of course 15 or more, 20 or more, 30 or more, 40 or more, 50 or more, 60 or more, 70 or more, 80 or more, 90 or more, 100 or more, 110 or more, 120 or more, 130 or more, 140 or more, or all the genes provided in Tables 1 and/or 2 below may be used. "Accession" as used in the context of the Tables herein as well as the present invention refers to the GenBank accession number of a sequence of each gene, the sequences of which are hereby incorporated by reference in their entireties as they are available from GenBank as accessed on the filing date of the present application. P value refers to values assigned as described in the Examples below. The indications of "E-xx" where "xx" is a two digit number refers to alternative notation for exponential figures where "E-xx" is "10^{-xx}". Thus in combination with the numbers to the left of "E-xx", the value being represented is the numbers to the left times 10^{-xx}. "Description" as used in the Tables provides a brief identifier of what the sequence/gene encodes.

Genes with a correlation identified by a p value below or about 0.02, below or about 0.01, below or about 0.005, or below or about 0.001 are preferred for use in the practice of the invention. The present invention includes the use of gene(s) the expression of which identify different ER+ breast cancer outcomes after TAM treatment to permit simultaneous identification of breast cancer survival outcome of a patient based upon assaying a breast cancer sample from said patient.

In a second aspect, the present invention relates to the identification and use of three sets of sequences for the determination of responsiveness to TAM treatment in ER+ breast cancer. The differential expression of these sequences in breast cancer relative to normal breast cells is used to predict TAM responsiveness in a subject. The identity of the sets of sequences were determined by use of ER+ primary breast cancers from 60 patients uniformly treated with adjuvant TAM. The cancers were analyzed using high-density oligonucleotide microarrays to identify gene expression patterns highly correlated with treatment outcome. Expression levels of IL17BR, CACNA1D, and HOXB13 were strongly predictive of clinical outcome. In contrast, a previously reported 70-gene prognosis signature was not a significant predictor of clinical outcome in these patients. Validation in an independent cohort of 31 TAM treated patients confirmed the predictive utility of these three genes.

In comparison with existing biomarkers, including ESR1, PGR, ERBB2 and EGFR, these genes are significantly more predictive of TAM response. Multivariate analysis indicated that these three genes were significant predictors of clinical outcome independent of tumor size, nodal status and tumor grade. TAM is the most effective systemic treatment for ER+ breast cancer. ER and progesterone receptor (PR) expression have been the major clinicopathological predictors for response to TAM. However, up to 40% of ER+ tumors fail to respond or develop resistance to TAM. The invention thus provides for the use of the identified biomarkers to allow better patient management by identifying patients who are more likely to benefit from TAM or other endocrine therapy and those who are likely to develop resistance and tumor recurrence.

As noted herein, the sequences(s) identified by the present invention are expressed in correlation with ER+ breast cells. For example, IL17RB, identified by I.M.A.G.E. Consortium Clusters NM_018725 and NM_172234 ("The I.M.A.G.E. Consortium: An Integrated Molecular Analysis of Genomes and their Expression," Lennon et al., 1996, Genomics 33:151-152; see also image.llnl.gov) has been found to be useful in predicting responsiveness to TAM treatment.

In preferred embodiments of the invention, any sequence, or unique portion thereof, of the IL17RB sequences of the cluster, as well as the UniGene *Homo sapiens* cluster Hs.5470, may be used. Similarly, any sequence encoding all or a part of the protein encoded by any IL17RB sequence disclosed herein may be used. Consensus sequences of I.M.A.G.E. Consortium clusters are as follows, with the assigned coding region (ending with a termination codon) underlined and preceded by the 5' untranslated and/or non-coding region and followed by the 3' untranslated and/or non-coding region:

SEQ ID NO:1 (consensus sequence for IL17RB, transcript variant 1, identified as NM_018725 or NM_018725.2)

agcgcagcgt gcgggtggcc tggatcccgc gcagtggccc ggcgatgtcg ctcgtgctgc
taagcctggc cgcgctgtgc aggagcgccg taccccgaga gccgaccgtt caatgtggct
ctgaaactgg gccatctcca gagtggatgc tacaacatga tctaatcccc ggagacttga
gggacctccg agtagaacct gttacaacta gtgttgcaac aggggactat tcaattttga
tgaatgtaag ctgggtactc cgggcagatg ccagcatccg cttgttgaag gccaccaaga
tttgtgtgac gggcaaaagc aacttccagt cctacagctg tgtgaggtgc aattacacag
aggccttcca gactcagacc agaccctctg gtggtaaatg gacattttcc tacatcggct

tccctgtaga gctgaacaca gtctatttca ttggggccca taatattcct aatgcaaata
tgaatgaaga tggcccttcc atgtctgtga atttcacctc accaggctgc ctagaccaca
taatgaaata taaaaaaaag tgtgtcaagg ccggaagcct gtgggatccg aacatcactg
cttgtaagaa gaatgaggag acagtagaag tgaacttcac aaccactccc ctgggaaaca
gatacatggc tcttatccaa cacagcacta tcatcgggtt ttctcaggtg tttgagccac
accagaagaa acaaacgcga gcttcagtgg tgattccagt gactggggat agtgaaggtg
ctacggtgca gctgactcca tattttccta cttgtggcag cgactgcatc cgacataaag
gaacagttgt gctctgccca caaacaggcg tccctttccc tctggataac aacaaaagca
agccgggagg ctggctgcct ctcctcctgc tgtctctgct ggtggccaca tgggtgctgg
tggcagggat ctatctaatg tggaggcacg aaaggatcaa gaagacttcc ttttctacca
ccacactact gccccccatt aaggttcttg tggtttaccc atctgaaata tgtttccatc
acacaatttg ttacttcact gaatttcttc aaaaccattg cagaagtgag gtcatccttg
aaaagtggca gaaaaagaaa atagcagaga tgggtccagt gcagtggctt gccactcaaa
agaaggcagc agacaaagtc gtcttccttc tttccaatga cgtcaacagt gtgtgcgatg
gtacctgtgg caagagcgag ggcagtccca gtgagaactc tcaagacctc ttccccttg
cctttaacct tttctgcagt gatctaagaa gccagattca tctgcacaaa tacgtggtgg
tctactttag agagattgat acaaaagacg attacaatgc tctcagtgtc tgccccaagt
accacctcat gaaggatgcc actgctttct gtgcagaact tctccatgtc aagcagcagg
tgtcagcagg aaaaagatca caagcctgcc acgatggctg ctgctccttg tagcccaccc
atgagaagca agagacctta aaggcttcct atcccaccaa ttacagggaa aaaacgtgtg
atgatcctga agcttactat gcagcctaca aacagcctta gtaattaaaa cattttatac
caataaaatt ttcaaatatt gctaactaat gtagcattaa ctaacgattg gaaactacat
ttacaacttc aaagctgttt tatacataga aatcaattac agttttaatt gaaaactata
accattttga taatgcaaca ataaagcatc ttcagccaaa catctagtct tccatagacc
atgcattgca gtgtacccag aactgtttag ctaatattct atgtttaatt aatgaatact
aactctaaga acccctcact gattcactca atagcatctt aagtgaaaaa ccttctatta
catgcaaaaa atcattgttt ttaagataac aaaagtaggg aataaacaag ctgaacccac
ttttaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa

SEQ ID NO:2 (consensus sequence for IL17RB, transcript variant 2, identified as NM_172234 or NM_172234.1)

agcgcagcgt gcgggtggcc tggatcccgc gcagtggccc ggcgatgtcg ctcgtgctgc
taagcctggc cgcgctgtgc aggagcgccg taccccgaga gccgaccgtt caatgtggct
ctgaaactgg gccatctcca gagtggatgc tacaacatga tctaatcccc ggagacttga
gggacctccg agtagaacct gttacaacta gtgttgcaac aggggactat tcaattttga
tgaatgtaag ctgggtactc cgggcagatg ccagcatccg cttgttgaag gccaccaaga
tttgtgtgac gggcaaaagc aacttccagt cctacagctg tgtgaggtgc aattacacag
aggccttcca gactcagacc agaccctctg gtggtaaatg gacattttcc tacatcggct
tccctgtaga gctgaacaca gtctatttca ttggggccca taatattcct aatgcaaata
tgaatgaaga tggcccttcc atgtctgtga atttcacctc accaggctgc ctagaccaca
taatgaaata taaaaaaaag tgtgtcaagg ccggaagcct gtgggatccg aacatcactg
cttgtaagaa gaatgaggag acagtagaag tgaacttcac aaccactccc ctgggaaaca
gatacatggc tcttatccaa cacagcacta tcatcgggtt ttctcaggtg tttgagccac

accagaagaa acaaacgcga gcttcagtgg tgattccagt gactggggat agtgaaggtg
ctacggtgca ggtaaagttc agtgagctgc tctggggagg gaagggacat agaagactgt
tccatcattc attgctttta aggatgagtt ctctcttgtc aaatgcactt ctgccagcag
acaccagtta agtggcgttc atgggggctc tttcgctgca gcctccaccg tgctgaggtc
aggaggccga cgtggcagtt gtggtccctt ttgcttgtat taatggctgc tgaccttcca
aagcactttt tattttcatt ttctgtcaca gacactcagg gatagcagta ccatttact
tccgcaagcc tttaactgca agatgaagct gcaaagggtt tgaaatggga aggtttgagt
tccaggcagc gtatgaactc tggagagggg ctgccagtcc tctctgggcc gcagcggacc
cagctggaac acaggaagtt ggagcagtag gtgctccttc acctctcagt atgtctcttt
caactctagt ttttgaggtg gggacacagg aggtccagtg ggacacagcc actccccaaa
gagtaaggag cttccatgct tcattccctg gcataaaaag tgctcaaaca caccagaggg
ggcaggcacc agccagggta tgatggctac taccctttc tggagaacca tagacttccc
ttactacagg gacttgcatg tcctaaagca ctggctgaag gaagccaaga ggatcactgc
tgctcctttt ttctagagga aatgtttgtc tacgtggtaa gatatgacct agccctttta
ggtaagcgaa ctggtatgtt agtaacgtgt acaaagttta ggttcagacc ccgggagtct
tgggcacgtg ggtctcgggt cactggtttt gactttaggg ctttgttaca gatgtgtgac
caaggggaaa atgtgcatga caacactaga ggtatgggcg aagccagaaa gaagggaagt
tttggctgaa gtaggagtct tggtgagatt ttgctctgat gcatggtgtg aactttctga
gcctcttgtt tttcctcagc tgactccata ttttcctact tgtggcagcg actgcatccg
acataaagga acagttgtgc tctgcccaca aacaggcgtc cctttccctc tggataacaa
caaaagcaag ccgggaggct ggctgcctct cctcctgctg tctctgctgg tggccacatg
ggtgctggtg gcagggatct atctaatgtg gaggcacgaa aggatcaaga agacttcctt
ttctaccacc acactactgc cccccattaa ggttcttgtg gtttacccat ctgaaatatg
tttccatcac acaatttgtt acttcactga atttcttcaa aaccattgca gaagtgaggt
catccttgaa aagtggcaga aaaagaaaat agcagagatg ggtccagtgc agtggcttgc
cactcaaaag aaggcagcag acaaagtcgt cttccttctt tccaatgacg tcaacagtgt
gtgcgatggt acctgtggca agagcgaggg cagtcccagt gagaactctc aagacctctt
ccccccttgcc tttaaccttt tctgcagtga tctaagaagc cagattcatc tgcacaaata
cgtggtggtc tactttagag agattgatac aaaagacgat tacaatgctc tcagtgtctg
ccccaagtac cacctcatga aggatgccac tgctttctgt gcagaacttc tccatgtcaa
gcagcaggtg tcagcaggaa aaagatcaca agcctgccac gatggctgct gctccttgta
gcccacccat gagaagcaag agaccttaaa ggcttcctat cccaccaatt acagggaaaa
aacgtgtgat gatcctgaag cttactatgc agcctacaaa cagccttagt aattaaaaca
ttttatacca ataaaatttt caaatattgc taactaatgt agcattaact aacgattgga
aactacattt acaacttcaa agctgtttta tacatagaaa tcaattacag ttttaattga
aaactataac cattttgata atgcaacaat aaagcatctt cagccaaaca tctagtcttc
catagaccat gcattgcagt gtacccagaa ctgtttagct aatattctat gtttaattaa
tgaatactaa ctctaagaac ccctcactga ttcactcaat agcatcttaa gtgaaaaacc
ttctattaca tgcaaaaaat cattgttttt aagataacaa aagtagggaa taaacaagct
gaacccactt ttaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa

I.M.A.G.E. Consortium Clone ID numbers and the corresponding GenBank accession numbers of sequences identified as belonging to the I.M.A.G.E. Consortium and UniGene clusters,

are listed below. Also included are sequences that are not identified as having a Clone ID number but still identitied as being those of IL17RB. The sequences include those of the "sense" and complementary strands sequences corresponding to IL17RB. The sequence of each GenBank accession number is presented in the attached Appendix.

5

Clone ID numbers	GenBank accession numbers
2985728	AW675096, AW673932, BC000980
5286745	BI602183
5278067	BI458542
5182255	BI823321
924000	AA514396
3566736	BF110326
3195409	BE466508
3576775	BF740045
2772915	AW299271
1368826	AA836217
1744837	AI203628
2285564	AI627783
2217709	AI744263
2103651	AI401622
2419487	AI826949
3125592	BE047352
2284721	AI911549
3643302	BF194822
1646910	AI034244
1647001	AI033911
3323709	BF064177
1419779	AA847767
2205190	AI538624
2295838	AI913613
2461335	AI942234
2130362	AI580483

2385555	AI831909
2283817	AI672344
2525596	AW025192
454687	AA677205
1285273	AA721647
3134106	BF115018
342259	W61238, W61239
1651991	AI032064
2687714	AW236941
3302808	BG057174
2544461	AW058532
122014	T98360, T98361
2139250	AI470845
2133899	AI497731
121300	T96629, T96740
162274	H25975, H25941
3446667	BE539514, BX282554
156864	R74038, R74129
4611491	BG433769
4697316	BG530489
429376	AA007528, AA007529
5112415	BI260259
701357	AA287951, AA287911
121909	T97852, T97745
268037	N40294
1307489	AA809841
1357543	AA832389
48442	H14692
1302619	AA732635
1562857	AA928257
1731938	AI184427
1896025	AI298577
2336350	AI692717

1520997	AA910922
240506	H90761
2258560	AI620122
1569921	AI793318, AA962325, AI733290
6064627	BQ226353
299018	W04890
5500181	BM455231
2484011	BI492426
4746376	BG674622
233783	BX111256
1569921	BX117618
450450	AA682806
1943085	AI202376
2250390	AI658949
4526156	BG403405
3249181	BE673417
2484395	AW021469
30515867	CF455736
2878155	AW339874
4556884	BG399724
3254505	BF475787
3650593	BF437145
233783	H64601
None (mRNA sequences)	AF212365, AF208110, AF208111, AF250309, AK095091
None	BM983744, CB305764, BM715988, BM670929, BI792416, BI715216, N56060, CB241389, AV660618, BX088671, CB154426, CA434589, CA412162, CA314073, BF921554, BF920093, AV685699, AV650175, BX483104, CD675121, BE081436, AW970151, AW837146, AW368264, D25960, AV709899, BX431018, AL535617, AL525465, BX453536, BX453537, AV728945, AV728939, AV727345

In one preferred embodiment, any sequence, or unique portion thereof, of the following IL17RB sequence, identified by AF208111 or AF208111.1, may be used in the practice of the invention.

SEQ ID NO:3 (sequence for IL17RB):

CGGCGATGTCGCTCGTGCTGATAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAG
AGCCGACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATG
ATCTAATCCCCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACTAGTGTTGCAA
CAGGGGACTATTCAATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCC
GCTTGTTGAAGGCCACCAAGATTTGTGTGACGGGCAAAAGCAACTTCCAGTCCTACAGCT
GTGTGAGGTGCAATTACACAGAGGCCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAAT
GGACATTTTCCTATATCGGCTTCCCTGTAGAGCTGAACACAGTCTATTTCATTGGGGCCC
ATAATATTCCTAATGCAAATATGAATGAAGATGGCCCTTCCATGTCTGTGAATTTCACCT
CACCAGGCTGCCTAGACCACATAATGAAATATAAAAAAAAAGTGTGTCAAGGCCGGAAGCC
TGTGGGATCCGAACATCACTGCTTGTAAGAAGAATGAGGAGACAGTAGAAGTGAACTTCA
CAACCACTCCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCACTATCATCGGGT
TTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGATTCCAG
TGACTGGGGATAGTGAAGGTGCTACGGTGCAGGTAAAGTTCAGTGAGCTGCTCTGGGGAG
GGAAGGGACATAGAAGACTGTTCCATCATTCATTGCTTTTAAGGATGAGTTCTCTCTTGT
CAAATGCACTTCTGCCAGCAGACACCAGTTAAGTGGCGTTCATGGGGGTTCTTTCGCTGC
AGCCTCCACCGTGCTGAGGTCAGGAGGCCGACGTGGCAGTTGTGGTCCCTTTTGCTTGTA
TTAATGGCTGCTGACCTTCCAAAGCACTTTTTATTTTCATTTTCTGTCACAGACACTCAG
GGATAGCAGTACCATTTTACTTCCGCAAGCCTTTAACTGCAAGATGAAGCTGCAAAGGGT
TTGAAATGGGAAGGTTTGAGTTCCAGGCAGCGTATGAACTCTGGAGAGGGGCTGCCAGTC
CTCTCTGGGCCGCAGCGGACCCAGCTGGAACACAGGAAGTTGGAGCAGTAGGTGCTCCTT
CACCTCTCAGTATGTCTCTTTCAACTCTAGTTTTTGAAGTGGGGACACAGGAAGTCCAGT
GGGGACACAGCCACTCCCCAAAGAATAAGGAACTTCCATGCTTCATTCCCTGGCATAAAA
AGTGNTCAAACACACCAGAGGGGGCAGGCACCAGCCAGGGTATGATGGGTACTACCCTTT
TCTGGAGAACCATAGACTTCCCTTACTACAGGGACTTGCATGTCCTAAAGCACTGGCTGA
AGGAAGCCAAGAGGATCACTGCTGCTCCTTTTTTGTAGAGGAAATGTTTGTGTACGTGGT
AAGATATGACCTAGCCCTTTTAGGTAAGCGAACTGGTATGTTAGTAACGTGTACAAAGTT
TAGGTTCAGACCCCGGGAGTCTTGGGCATGTGGGTCTCGGGTCACTGGTTTTGACTTTAG
GGCTTTGTTACAGATGTGTGACCAAGGGGAAAATGTGCATGACAACACTAGAGGTAGGGG
CGAAGCCAGAAAGAAGGGAAGTTTTGGCTGAAGTAGGAGTCTTGGTGAGATTTTGCTGTG
ATGCATGGTGTGAACTTTCTGAGCCTCTTGTTTTTCCTCAGCTGACTCCATATTTTCCTA
CTTGTGGCAGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCG
TCCCTTTCCCTCTGGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGC
TGTCTCTGCTGGTGGCCACATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACG
AAAGGATCAAGAAGACTTCCTTTTCTACCACCACACTACTGCCCCCCATTAAGGTTCTTG
TGGTTTACCCATCTGAAATATGTTTCCATCACACAATTTGTTACTTCACTGAATTTCTTC
AAAACCATTGCAGAAGTGAGGTCATCCTTGAAAAGTGGCAGAAAAAGAAAATAGCAGAGA

TGGGTCCAGTGCAGTGGCTTGCCACTCAAAAGAAGGCAGCAGACAAAGTCGTCTTCCTTC
TTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCCA
GTGAGAACTCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAA
GCCAGATTCATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACG
ATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACTTCATGAAGGATGCCACTGCTTTCT
GTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCC
ACGATGGCTGCTGCTCCTTGTAGCCCACCCATGAGAAGCAAGAGACCTTAAAGGCTTCCT
ATCCCACCAATTACAGGGAAAAAACGTGTGATGATCCTGAAGCTTACTATGCAGCCTACA
AACAGCCTTAGTAATTAAAACATTTTATACCAATAAAATTTTCAAATATTACTAACTAAT
GTAGCATTAACTAACGATTGGAAACTACATTTACAACTTCAAAGCTGTTTTATACATAGA
AATCAATTACAGCTTTAATTGAAAACTGTAACCATTTTGATAATGCAACAATAAAGCATC
TTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

In another set of preferred embodiments of the invention, any sequence, or unique portion thereof, of the CACNA1D sequences of the I.M.A.G.E. Consortium cluster NM_000720, as well as the UniGene Homo sapiens cluster Hs.399966, may be used. Similarly, any sequence encoding all or a part of the protein encoded by any CACNA1D sequence disclosed herein may be used. The consensus sequence of the I.M.A.G.E. Consortium cluster is as follows, with the assigned coding region (ending with a termination codon) underlined and preceded by the 5' untranslated and/or non-coding region and followed by the 3' untranslated and/or non-coding region:

SEQ ID NO:4 (consensus sequence for CACNA1D, identified as NM_000720 or NM_000720.1)

agaataaggg cagggaccgc ggctcctatc tcttggtgat ccccttcccc attccgcccc
cgcctcaacg cccagcacag tgccctgcac acagtagtcg ctcaataaat gttcgtggat
gatgatgatg atgatgatga aaaaaatgca gcatcaacgg cagcagcaag cggaccacgc
gaacgaggca aactatgcaa gaggcaccag acttcctctt tctggtgaag gaccaacttc
tcagccgaat agctccaagc aaactgtcct gtcttggcaa gctgcaatcg atgctgctag
acaggccaag gctgcccaaa ctatgagcac ctctgcaccc ccacctgtag gatctctctc
ccaaagaaaa cgtcagcaat acgccaagag caaaaaacag ggtaactcgt ccaacagccg
acctgcccgc gccctttttct gtttatcact caataacccc atccgaagag cctgcattag
tatagtggaa tggaaaccat ttgacatatt tatattattg gctatttttg ccaattgtgt
ggccttagct atttacatcc cattccctga agatgattct aattcaacaa atcataactt
ggaaaaagta gaatatgcct tcctgattat ttttacagtc gagacatttt tgaagattat
agcgtatgga ttattgctac atcctaatgc ttatgttagg aatggatgga atttactgga
ttttgttata gtaatagtag gattgtttag tgtaattttg gaacaattaa ccaaagaaac
agaaggcggg aaccactcaa gcggcaaatc tggaggcttt gatgtcaaag ccctccgtgc
ctttcgagtg ttgcgaccac ttcgactagt gtcaggggtg cccagtttac aagttgtcct
gaactccatt ataaaagcca tggttcccct ccttcacata gccctttttgg tattatttgt

aatcataatc tatgctatta taggattgga acttttatt ggaaaaatgc acaaaacatg
ttttttgct gactcagata tcgtagctga agaggaccca gctccatgtg cgttctcagg
gaatggacgc cagtgtactg ccaatggcac ggaatgtagg agtggctggg ttggcccgaa
cggaggcatc accaactttg ataactttgc ctttgccatg cttactgtgt ttcagtgcat
caccatggag ggctggacag acgtgctcta ctgggtaaat gatgcgatag gatgggaatg
gccatgggtg tattttgtta gtctgatcat ccttggctca tttttcgtcc ttaacctggt
tcttggtgtc cttagtggag aattctcaaa ggaaagagag aaggcaaaag cacggggaga
tttccagaag ctccgggaga agcagcagct ggaggaggat ctaaagggct acttggattg
gatcacccaa gctgaggaca tcgatccgga gaatgaggaa gaaggaggag aggaaggcaa
acgaaatact agcatgccca ccagcgagac tgagtctgtg aacacagaga acgtcagcgg
tgaaggcgag aaccgaggct gctgtggaag tctctggtgc tggtggagac ggagaggcgc
ggccaaggcg gggccctctg ggtgtcggcg gtggggtcaa gccatctcaa aatccaaact
cagccgacgc tggcgtcgct ggaaccgatt caatcgcaga agatgtaggg ccgccgtgaa
gtctgtcacg ttttactggc tggttatcgt cctggtgttt ctgaacacct taaccatttc
ctctgagcac tacaatcagc cagattggtt gacacagatt caagatattg ccaacaaagt
cctcttggct ctgttcacct gcgagatgct ggtaaaaatg tacagcttgg gcctccaagc
atatttcgtc tctcttttca accggtttga ttgcttcgtg gtgtgtggtg gaatcactga
gacgatcctg gtggaactgg aaatcatgtc tcccctgggg atctctgtgt ttcggtgtgt
gcgcctctta agaatcttca aagtgaccag gcactggact tccctgagca acttagtggc
atccttatta aactccatga agtccatcgc ttcgctgttg cttctgcttt ttctcttcat
tatcatcttt tccttgcttg ggatgcagct gtttggcggc aagtttaatt ttgatgaaac
gcaaaccaag cggagcacct ttgacaattt ccctcaagca cttctcacag tgttccagat
cctgacaggc gaagactgga atgctgtgat gtacgatggc atcatggctt acgggggccc
atcctcttca ggaatgatcg tctgcatcta cttcatcatc ctcttcattt gtggtaacta
tattctactg aatgtcttct tggccatcgc tgtagacaat ttggctgatg ctgaaagtct
gaacactgct cagaaagaag aagcggaaga aaaggagagg aaaaagattg ccagaaaaga
gagcctagaa aataaaaaga acaacaaacc agaagtcaac cagatagcca acagtgacaa
caaggttaca attgatgact atagagaaga ggatgaagac aaggacccct atccgccttg
cgatgtgcca gtaggggaag aggaagagga agaggaggag gatgaacctg aggttcctgc
cggaccccgt cctcgaagga tctcggagtt gaacatgaag gaaaaaattg cccccatccc
tgaagggagc gctttcttca ttcttagcaa gaccaacccg atccgcgtag gctgccacaa
gctcatcaac caccacatct tcaccaacct catccttgtc ttcatcatgc tgagcagcgc
tgccctggcc gcagaggacc ccatccgcag ccactccttc cggaacacga tactgggtta
ctttgactat gccttcacag ccatctttac tgttgagatc ctgttgaaga tgacaacttt
tggagctttc ctccacaaag gggccttctg caggaactac ttcaatttgc tggatatgct
ggtggttggg gtgtctctgg tgtcatttgg gattcaatcc agtgccatct ccgttgtgaa
gattctgagg gtcttaaggg tcctgcgtcc cctcagggcc atcaacagag caaaaggact
taagcacgtg gtccagtgcg tcttcgtggc catccggacc atcggcaaca tcatgatcgt
cactaccctc ctgcagttca tgtttgcctg tatcggggtc cagttgttca aggggaagtt
ctatcgctgt acggatgaag ccaaaagtaa ccctgaagaa tgcaggggac ttttcatcct
ctacaaggat ggggatgttg acagtcctgt ggtccgtgaa cggatctggc aaaacagtga
tttcaacttc gacaacgtcc tctctgctat gatggcgctc ttcacagtct ccacgtttga
gggctggcct gcgttgctgt ataaagccat cgactcgaat ggagagaaca tcggcccaat
ctacaaccac cgcgtggaga tctccatctt cttcatcatc tacatcatca ttgtagcttt
cttcatgatg aacatctttg tgggctttgt catcgttaca tttcaggaac aaggagaaaa

agagtataag aactgtgagc tggacaaaaa tcagcgtcag tgtgttgaat acgccttgaa
agcacgtccc ttgcggagat acatccccaa aaacccctac cagtacaagt tctggtacgt
ggtgaactct tcgcctttcg aatacatgat gtttgtcctc atcatgctca acacactctg
cttggccatg cagcactacg agcagtccaa gatgttcaat gatgccatgg acattctgaa
catggtcttc accggggtgt tcaccgtcga gatggttttg aaagtcatcg catttaagcc
taaggggtat tttagtgacg cctggaacac gtttgactcc ctcatcgtaa tcggcagcat
tatagacgtg gccctcagcg aagcggaccc aactgaaagt gaaaatgtcc ctgtcccaac
tgctacacct gggaactctg aagagagcaa tagaatctcc atcacctttt tccgtctttt
ccgagtgatg cgattggtga agcttctcag caggggggaa ggcatccgga cattgctgtg
gacttttatt aagtcctttc aggcgctccc gtatgtggcc ctcctcatag ccatgctgtt
cttcatctat gcggtcattg gcatgcagat gtttgggaaa gttgccatga gagataacaa
ccagatcaat aggaacaata acttccagac gtttccccag gcggtgctgc tgctcttcag
gtgtgcaaca ggtgaggcct ggcaggagat catgctggcc tgtctcccag ggaagctctg
tgaccctgag tcagattaca accccgggga ggagtataca tgtgggagca actttgccat
tgtctatttc atcagttttt acatgctctg tgcatttctg atcatcaatc tgtttgtggc
tgtcatcatg gataatttcg actatctgac ccgggactgg tctattttgg ggcctcacca
tttagatgaa ttcaaaagaa tatggtcaga atatgaccct gaggcaaagg gaaggataaa
acaccttgat gtggtcactc tgcttcgacg catccagcct cccctggggt ttgggaagtt
atgtccacac agggtagcgt gcaagagatt agttgccatg aacatgcctc tcaacagtga
cgggacagtc atgtttaatg caaccctgtt tgctttggtt cgaacggctc ttaagatcaa
gaccgaaggg aacctggagc aagctaatga agaacttcgg gctgtgataa agaaaatttg
gaagaaaacc agcatgaaat tacttgacca agttgtccct ccagctggtg atgatgaggt
aaccgtgggg aagttctatg ccactttcct gatacaggac tactttagga aattcaagaa
acggaaagaa caaggactgg tgggaaagta ccctgcgaag aacaccacaa ttgccctaca
ggcgggatta aggacactgc atgacattgg gccagaaatc cggcgtgcta tatcgtgtga
tttgcaagat gacgagcctg aggaaacaaa acgagaagaa gaagatgatg tgttcaaaag
aaatggtgcc ctgcttggaa accatgtcaa tcatgttaat agtgatagga gagattccct
tcagcagacc aataccaccc accgtcccct gcatgtccaa aggccttcaa ttccacctgc
aagtgatact gagaaaccgc tgtttcctcc agcaggaaat tcggtgtgtc ataaccatca
taaccataat tccataggaa agcaagttcc cacctcaaca aatgccaatc tcaataatgc
caatatgtcc aaagctgccc atggaaagcg gcccagcatt gggaaccttg agcatgtgtc
tgaaaatggg catcattctt cccacaagca tgaccgggag cctcagagaa ggtccagtgt
gaaaagaacc cgctattatg aaacttacat taggtccgac tcaggagatg aacagctccc
aactatttgc cgggaagacc cagagataca tggctatttc agggaccccc actgcttggg
ggagcaggag tatttcagta gtgaggaatg ctacgaggat gacagctcgc ccacctggag
caggcaaaac tatggctact acagcagata cccaggcaga aacatcgact ctgagaggcc
ccgaggctac catcatcccc aaggattctt ggaggacgat gactcgcccg tttgctatga
ttcacggaga tctccaagga gacgcctact acctcccacc ccagcatccc accggagatc
ctccttcaac tttgagtgcc tgcgccggca gagcagccag gaagaggtcc cgtcgtctcc
catcttcccc catcgcacgg ccctgcctct gcatctaatg cagcaacaga tcatggcagt
tgccggccta gattcaagta aagcccagaa gtactcaccg agtcactcga cccggtcgtg
ggccaccccc ccagcaaccc ctccctaccg ggactggaca ccgtgctaca ccccccctgat
ccaagtggag cagtcagagg ccctggacca ggtgaacggc agcctgccgt ccctgcaccg
cagctcctgg tacacagacg agcccgacat ctcctaccgg actttcacac cagccagcct
gactgtcccc agcagcttcc ggaacaaaaa cagcgacaag cagaggagtg cggacagctt

ggtggaggca gtcctgatat ccgaaggctt gggacgctat gcaagggacc caaaatttgt
gtcagcaaca aaacacgaaa tcgctgatgc ctgtgacctc accatcgacg agatggagag
tgcagccagc accctgctta atgggaacgt gcgtccccga gccaacgggg atgtgggccc
cctctcacac cggcaggact atgagctaca ggactttggt cctggctaca gcgacgaaga
gccagaccct gggagggatg aggaggacct ggcggatgaa atgatatgca tcaccacctt
gtagccccca gcgaggggca gactggctct ggcctcaggt ggggcgcagg agagccaggg
gaaaagtgcc tcatagttag gaaagtttag gcactagttg ggagtaatat tcaattaatt
agacttttgt ataagagatg tcatgcctca agaaagccat aaacctggta ggaacaggtc
ccaagcggtt gagcctggca gagtaccatg cgctcggccc cagctgcagg aaacagcagg
ccccgccctc tcacagagga tgggtgagga ggccagacct gccctgcccc attgtccaga
tgggcactgc tgtggagtct gcttctccca tgtaccaggg caccaggccc acccaactga
aggcatggcg gcggggtgca ggggaaagtt aaaggtgatg acgatcatca cacctcgtgt
cgttacctca gccatcggtc tagcatatca gtcactgggc ccaacatatc catttttaaa
ccctttcccc caaatacact gcgtcctggt tcctgtttag ctgttctgaa ata

I.M.A.G.E. Consortium Clone ID numbers and the corresponding GenBank accession numbers of sequences identified as belonging to the I.M.A.G.E. Consortium and UniGene clusters, are listed below. Also included are sequences that are not identified as having a Clone ID number but still identitied as being those of CACNA1D. The sequences include those of the "sense" and complementary strands sequences corresponding to CACNA1D. The sequence of each GenBank accession number is presented in the attached Appendix.

Clone ID numbers	GenBank accession numbers
5676430	BM128550
5197948	BI755471
6027638	BQ549084, BQ549571
2338956	AI693324
36581	R25307, R46658
49630	H29256, H29339
4798765	BG716371
2187310	AI537488
838231	AA458692
2111614	AI393327
2183482	AI520947
1851007	AI248998

1675503	AI075844
2434923	AI869807
2434924	AI869800
1845827	AI243110
2511756	AI955764
628568	AA192669, AA192157
2019331	AI361691
2337381	AI914244
2503579	AW008769
2503626	AW008794
1160989	AA877582
1653475	AI051972
1627755	AI017959
287750	N79331, N62240
1867677	AI240933
1618303	AI015031
1881344	AI290994
1408031	AA861160
1557035	AA915941
956303	AA493341
2148234	AI467998
1499899	AA885585
1647592	AI033648
2341185	AI697633
981603	AA523647
6281678	BQ710377
6278348	BQ706920
5876024	BQ016847
6608849	CA943595
5440464	BM008196
5209489	BI769856
5183025	BI758971
880540	AA468565

757337	AA437099
6608849	CA867864
461797	AA682690
434787	AA701888
6151588	BU182632
6295618	BQ898429
6300779	BQ711800
434811	AA703120
1568025	AA978315
3220210	BE550599
3214121	BE502741
3009312	AW872382
2733394	AW444663
2872156	AW341279
30514550	CF456750
2718456	AW139850
2543682	AW029633
2492730	AI963788
2545866	AI951788
2272081	AI680744
2152336	AI601252
2146429	AI459166
1274498	AA885750
2272081	BX092736
287750	BX114568
3233645	BE672659
289209	N78509, N73668
277086	N46744, N39597
3272340	BF439267
3273859	BF436153
3568401	BF110611
None (mRNA sequences)	M76558, AF088004, M83566
None	CB410657, BQ372430, BQ366601, BQ324528,

	BQ318830, AL708030, BM509161, N85902, BQ774355, CA774243, CA436347, CA389011, BU679327, BU608029, BU073743, BE175413, AW969248, AI908115, BF754485, BI015409, BG202552, BF883669, BF817590, BF807128, BF806160, BF805244, BF805235, BF805080, T27949, BE836638, BE770685, BE769065,

In one preferred embodiment, any sequence, or unique portion thereof, of the following CACNA1D sequence, identified by AI240933 or AI240933.1, may be used in the practice of the invention.

SEQ ID NO:5 (sequence for CACNA1D):

TTTTTTTTTTTTTTTTTTTTTCTTACAAAGAAAAATTTAATATTCGATGAGAGGTTGAAC
CAGGCTTAAAGCAGACATACTAGGAAATGGTGCAGCCTGTAAGAATGCCAGTTTGTAAGT
ACTGACTTTGGAAAAGATCATCGCCTCTATCAGACACTTAGGGTCCTGGTCTGGCAATTT
TGGCCTGATGTGATGCCACAAGACCCAACAGAGAGAGACACAGAGTCCAGGATAATGTTG
ACAGTGGTGTAGCCCTTTAGGAGAAATGGCGCTCCCTGCGGCTGGTATTAGGTTACCATT
GGCACCGAAGGAACCAGGAGGATAAGAATATCCATAATTTCAGAGCTGCCCTGGCACAGT
ACCTGCCCCGTCGGAGGCTCTCACTGGCAAATGACAGCTCTGTGCAAGGAGCACTCCCAA
GTATAAAAATTATTACACAGTTTTATTCTGAAGAACATTTTGCATTTTAATAAAAAAGGA
TTTATGTCAGGAAAGAGTCATTTACAAACCTTGAAGTGTTTTTGCCTGGATCAGAGTAAG
AATGTCTTAAGAAGAGGTTTGTAAGGTCTTCATAACAAAGTGGTGTTTGTTATTTACAAA
AAAAAAAAAAAAAAAAATTAACAGGTTGTCTGTATACTATTAAAAATTTTGGACCAAAAA
AAAAAAAAAAAAAA

In another set of preferred embodiments of the invention, any sequence, or unique portion thereof, of the HOXB13 sequences of the I.M.A.G.E. Consortium cluster NM_006361, as well as the UniGene Homo sapiens cluster Hs.66731, may be used. Similarly, any sequence encoding all or a part of the protein encoded by any HOXB13 sequence disclosed herein may be used. The consensus sequence of the I.M.A.G.E. Consortium cluster is as follows, with the assigned coding region (ending with a termination codon) underlined and preceded by the 5' untranslated and/or non-coding region and followed by the 3' untranslated and/or non-coding region:

SEQ ID NO:6 (consensus sequence for HOXB13, identified as NM_006361or NM_006361.2)

cgaatgcagg cgacttgcga gctgggagcg atttaaaacg ctttggattc ccccggcctg
ggtggggaga gcgagctggg tgccccctag attccccgcc cccgcacctc atgagccgac
cctcggctcc atggagcccg gcaattatgc caccttggat ggagccaagg atatcgaagg
cttgctggga gcgggagggg ggcggaatct ggtcgcccac tcccctctga ccagccaccc
agcggcgcct acgctgatgc ctgctgtcaa ctatgccccc ttggatctgc caggctcggc
ggagccgcca aagcaatgcc acccatgccc tggggtgccc caggggacgt ccccagctcc
cgtgccttat ggttactttg gaggcgggta ctactcctgc cgagtgtccc ggagctcgct
gaaaccctgt gcccaggcag ccaccctggc cgcgtacccc gcggagactc ccacggccgg
ggaagagtac cccagtcgcc ccactgagtt tgccttctat ccgggatatc cgggaaccta
ccacgctatg gccagttacc tggacgtgtc tgtggtgcag actctgggtg ctcctggaga
accgcgacat gactccctgt tgcctgtgga cagttaccag tcttgggctc tcgctggtgg
ctggaacagc cagatgtgtt gccagggaga acagaaccca ccaggtccct tttggaaggc
agcatttgca gactccagcg ggcagcaccc tcctgacgcc tgcgcctttc gtcgcggccg
caagaaacgc attccgtaca gcaaggggca gttgcgggag ctggagcggg agtatgcggc
taacaagttc atcaccaagg acaagaggcg caagatctcg gcagccacca gcctctcgga
gcgccagatt accatctggt ttcagaaccg ccgggtcaaa gagaagaagg ttctcgccaa
ggtgaagaac agcgctaccc cttaagagat ctccttgcct gggtgggagg agcgaaagtg
ggggtgtcct ggggagacca gaaacctgcc aagcccaggc tggggccaag gactctgctg
agaggcccct agagacaaca cccttcccag gccactggct gctggactgt tcctcaggag
cggcctgggt acccagtatg tgcagggaga cggaacccca tgtgacaggc ccactccacc
agggttccca aagaacctgg cccagtcata atcattcatc ctcacagtgg caataatcac
gataaccagt

I.M.A.G.E. Consortium Clone ID numbers and the corresponding GenBank accession numbers of sequences identified as belonging to the I.M.A.G.E. Consortium and UniGene clusters, are listed below. Also included are sequences that are not identified as having a Clone ID number but still identitied as being those of HOXB13. The sequences include those of the "sense" and complementary strands sequences corresponding to HOXB13. The sequence of each GenBank accession number is presented in the attached Appendix.

Clone ID numbers	GenBank accession numbers
4250486	BF676461, BC007092
5518335	BM462617
4874541	BG752489
4806039	BG778198
3272315	CB050884, CB050885

4356740	BF965191
6668163	BU930208
1218366	AA807966
2437746	AI884491
1187697	AA652388
3647557	BF446158
1207949	AA657924
1047774	AA644637
3649397	BF222357
971664	AA527613
996191	AA533227
813481	AA456069, AA455572, BX117624
6256333	BQ673782
2408470	AI814453
2114743	AI417272
998548	AA535663
2116027	AI400493
3040843	AW779219
1101311	AA594847
1752062	AI150430
898712	AA494387
1218874	AA662643
2460189	AI935940
986283	AA532530
1435135	AA857572
1871750	AI261980
3915135	BE888751
2069668	AI378797
667188	AA234220, AA236353
1101561	AA588193
1170268	AI821103, AI821851, AA635855
2095067	AI420753
4432770	BG180547

783296	AA468306, AA468232
3271646	CB050115, CB050116
1219276	AA661819
30570598	CF146837
30570517	CF146763
30568921	CF144902
3099071	CF141511
3096992	CF139563
3096870	CF139372
3096623	CF139319
3096798	CF139275
30572408	CF122893
2490082	AI972423
2251055	AI918975
2419308	AI826991
2249105	AI686312
2243362	AI655923
30570697	CF146922
3255712	BF476369
3478356	BF057410
3287977	BE645544
3287746	BE645408
3621499	BE388501
30571128	CF147366
30570954	CF147143
None (mRNA sequences)	BT007410, BC007092, U57052, U81599
None	CB120119, CB125764, AU098628, CB126130, BI023924, BM767063, BM794275, BQ363211, BM932052, AA357646, AW609525, CB126919, AW609336, AW609244, BF855145, AU126914, CB126449, AW582404, BX641644

In one preferred embodiment, any sequence, or unique portion thereof, of the following HOXB13 sequence, identified by BC007092 or BC007092.1, may be used in the practice of the invention.

SEQ ID NO:7 (sequence for HOXB13):

GGATTCCCCCGGCCTGGGTGGGGAGAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCG
CACCTCATGAGCCGACCCTCGGCTCCATGGAGCCCGGCAATTATGCCACCTTGGATGGAG
CCAAGGATATCGAAGGCTTGCTGGGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCC
CTCTGACCAGCCACCCAGCGGCGCCTACGCTGATGCCTGCTGTCAACTATGCCCCCTTGG
ATCTGCCAGGCTCGGCGGAGCCGCCAAAGCAATGCCACCCATGCCCTGGGGTGCCCCAGG
GGACGTCCCCAGCTCCCGTGCCTTATGGTTACTTTGGAGGCGGGTACTACTCCTGCCGAG
TGTCCCGGAGCTCGCTGAAACCCTGTGCCCAGGCAGCCACCCTGGCCGCGTACCCCGCGG
AGACTCCCACGGCCGGGGAAGAGTACCCCAGCCGCCCCACTGAGTTTGCCTTCTATCCGG
GATATCCGGGAACCTACCAGCCTATGGCCAGTTACCTGGACGTGTCTGTGGTGCAGACTC
TGGGTGCTCCTGGAGAACCGCGACATGACTCCCTGTTGCCTGTGGACAGTTACCAGTCTT
GGGCTCTCGCTGGTGGCTGGAACAGCCAGATGTGTTGCCAGGGAGAACAGAACCCACCAG
GTCCCTTTTGGAAGGCAGCATTTGCAGACTCCAGCGGGCAGCACCCTCCTGACGCCTGCG
CCTTTCGTCGCGGCCGCAAGAAACGCATTCCGTACAGCAAGGGGCAGTTGCGGGAGCTGG
AGCGGGAGTATGCGGCTAACAAGTTCATCACCAAGGACAAGAGGCGCAAGATCTCGGCAG
CCACCAGCCTCTCGGAGCGCCAGATTACCATCTGGTTTCAGAACCGCCGGGTCAAAGAGA
AGAAGGTTCTCGCCAAGGTGAAGAACAGCGCTACCCCTTAAGAGATCTCCTTGCCTGGGT
GGGAGGAGCGAAAGTGGGGGTGTCCTGGGGAGACCAGGAACCTGCCAAGCCCAGGCTGGG
GCCAAGGACTCTGCTGAGAGGCCCCTAGAGACAACACCCTTCCCAGGCCACTGGCTGCTG
GACTGTTCCTCAGGAGCGGCCTGGGTACCCAGTATGTGCAGGGAGACGGAACCCCATGTG
ACAGCCCACTCCACCAGGGTTCCCAAAGAACCTGGCCCAGTCATAATCATTCATCCTGAC
AGTGGCAATAATCACGATAACCAGTACTAGCTGCCATGATCGTTAGCCTCATATTTTCTA
TCTAGAGCTCTGTAGAGCACTTTAGAAACCGCTTTCATGAATTGAGCTAATTATGAATAA
ATTTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Sequences identified by SEQ ID NO. are provided using conventional representations of a DNA strand starting from the 5' phosphate linked end to the 3' hydroxyl linked end. The assignment of coding regions is generally by comparison to available consensus sequence(s) and therefore may contain inconsistencies relative to other sequences assigned to the same cluster. These have no effect on the practice of the invention because the invention can be practiced by use of shorter segments (or combinations thereof) of sequences unique to each of the three sets described above and not affected by inconsistencies. As non-limiting examples, a segment of IL17BR, CACNA1D, or HOXB13 nucleic acid sequence composed of a 3' untranslated region

sequence and/or a sequence from the 3' end of the coding region may be used as a probe for the detection of IL17BR, CACNA1D, or HOXB13 expression, respectively, without being affected by the presence of any inconsistency in the coding regions due to differences between sequences. Similarly, the use of an antibody which specifically recognizes IL17BR, CACNA1D, or HOXB13 protein to detect its expression would not be affected by the presence of any inconsistency in the representation of the coding regions provided above.

As will be appreciated by those skilled in the art, some of the above sequences include 3' poly A (or poly T on the complementary strand) stretches that do not contribute to the uniqueness of the disclosed sequences. The invention may thus be practiced with sequences lacking the 3' poly A (or poly T) stretches. The uniqueness of the disclosed sequences refers to the portions or entireties of the sequences which are found only in IL17BR, CACNA1D, or HOXB13 nucleic acids, including unique sequences found at the 3' untranslated portion of the genes. Preferred unique sequences for the practice of the invention are those which contribute to the consensus sequences for each of the three sets such that the unique sequences will be useful in detecting expression in a variety of individuals rather than being specific for a polymorphism present in some individuals. Alternatively, sequences unique to an individual or a subpopulation may be used. The preferred unique sequences are preferably of the lengths of polynucleotides of the invention as discussed herein.

To determine the (increased or decreased) expression levels of the above described sequences in the practice of the present invention, any method known in the art may be utilized. In one preferred embodiment of the invention, expression based on detection of RNA which hybridizes to polynucleotides containing the above described sequences is used. This is readily performed by any RNA detection or amplification+detection method known or recognized as equivalent in the art such as, but not limited to, reverse transcription-PCR (optionally real-time PCR), the methods disclosed in U.S. Patent Application 10/062,857 entitled "Nucleic Acid Amplification" filed on October 25, 2001 as well as U.S. Provisional Patent Applications 60/298,847 (filed June 15, 2001) and 60/257,801 (filed December 22, 2000), the methods disclosed in U.S. Patent 6,291,170, and quantitative PCR. Methods to identify increased RNA stability (resulting in an observation of increased expression) or decreased RNA stability (resulting in an

observation of decreased expression) may also be used. These methods include the detection of sequences that increase or decrease the stability of mRNAs containing the IL17BR, CACNA1D, or HOXB13 sequences disclosed herein. These methods also include the detection of increased mRNA degradation.

In particularly preferred embodiments of the invention, polynucleotides having sequences present in the 3' untranslated and/or non-coding regions of the above disclosed sequences are used to detect expression or non-expression of IL17BR, CACNA1D, or HOXB13 sequences in breast cells in the practice of the invention. Such polynucleotides may optionally contain sequences found in the 3' portions of the coding regions of the above disclosed sequences. Polynucleotides containing a combination of sequences from the coding and 3' non-coding regions preferably have the sequences arranged contiguously, with no intervening heterologous sequence(s).

Alternatively, the invention may be practiced with polynucleotides having sequences present in the 5' untranslated and/or non-coding regions of IL17BR, CACNA1D, or HOXB13 sequences in breast cells to detect their levels of expression. Such polynucleotides may optionally contain sequences found in the 5' portions of the coding regions. Polynucleotides containing a combination of sequences from the coding and 5' non-coding regions preferably have the sequences arranged contiguously, with no intervening heterologous sequence(s). The invention may also be practiced with sequences present in the coding regions of IL17BR, CACNA1D, or HOXB13.

Preferred polynucleotides contain sequences from 3' or 5' untranslated and/or non-coding regions of at least about 20, at least about 22, at least about 24, at least about 26, at least about 28, at least about 30, at least about 32, at least about 34, at least about 36, at least about 38, at least about 40, at least about 42, at least about 44, or at least about 46 consecutive nucleotides. The term "about" as used in the previous sentence refers to an increase or decrease of 1 from the stated numerical value. Even more preferred are polynucleotides containing sequences of at least or about 50, at least or about 100, at least about or 150, at least or about 200, at least or about 250, at least or about 300, at least or about 350, or at least or about 400 consecutive nucleotides. The term "about" as used in the preceding sentence refers to an increase or decrease of 10% from the stated numerical value.

Sequences from the 3' or 5' end of the above described coding regions as found in polynucleotides of the invention are of the same lengths as those described above, except that they would naturally be limited by the length of the coding region. The 3' end of a coding region may include sequences up to the 3' half of the coding region. Conversely, the 5' end of a coding region may include sequences up the 5' half of the coding region. Of course the above described sequences, or the coding regions and polynucleotides containing portions thereof, may be used in their entireties.

Polynucleotides combining the sequences from a 3' untranslated and/or non-coding region and the associated 3' end of the coding region are preferably at least or about 100, at least about or 150, at least or about 200, at least or about 250, at least or about 300, at least or about 350, or at least or about 400 consecutive nucleotides. Preferably, the polynucleotides used are from the 3' end of the gene, such as within about 350, about 300, about 250, about 200, about 150, about 100, or about 50 nucleotides from the polyadenylation signal or polyadenylation site of a gene or expressed sequence. Polynucleotides containing mutations relative to the sequences of the disclosed genes may also be used so long as the presence of the mutations still allows hybridization to produce a detectable signal.

In another embodiment of the invention, polynucleotides containing deletions of nucleotides from the 5' and/or 3' end of the above disclosed sequences may be used. The deletions are preferably of 1-5, 5-10, 10-15, 15-20, 20-25, 25-30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-125, 125-150, 150-175, or 175-200 nucleotides from the 5' and/or 3' end, although the extent of the deletions would naturally be limited by the length of the disclosed sequences and the need to be able to use the polynucleotides for the detection of expression levels.

Other polynucleotides of the invention from the 3' end of the above disclosed sequences include those of primers and optional probes for quantitative PCR. Preferably, the primers and probes are those which amplify a region less than about 350, less than about 300, less than about 250, less than about 200, less than about 150, less than about 100, or less than about 50 nucleotides from the from the polyadenylation signal or polyadenylation site of a gene or expressed sequence.

In yet another embodiment of the invention, polynucleotides containing portions of the above disclosed sequences including the 3' end may be used in the practice of the invention. Such

polynucleotides would contain at least or about 50, at least or about 100, at least about or 150, at least or about 200, at least or about 250, at least or about 300, at least or about 350, or at least or about 400 consecutive nucleotides from the 3' end of the disclosed sequences.

The invention thus also includes polynucleotides used to detect IL17BR, CACNA1D, or HOXB13 expression in breast cells. The polynucleotides may comprise a shorter polynucleotide consisting of sequences found in the above provided SEQ ID NOS in combination with heterologous sequences not naturally found in combination with IL17BR, CACNA1D, or HOXB13 sequences.

As non-limiting examples, a polynucleotide comprising one of the following sequences may be used in the practice of the invention.

SEQ ID NO:8:
CAATTACAGGGAAAAAACGTGTGATGATCCTGAAGCTTACTATGCAGCCTACAAACAGCC

SEQ ID NO:9:
GCTCTCACTGGCAAATGACAGCTCTGTGCAAGGAGCACTCCCAAGTATAAAAATTATTAC

SEQ ID NO:10:
GATCGTTAGCCTCATATTTTCTATCTAGAGCTCTGTAGAGCACTTTAGAAACCGCTTTCA

Stated differently, the invention may be practiced with a polynucleotide consisting of the sequence of SEQ ID NOS:8, 9 or 10 in combination with one or more heterologous sequences that are not normally found with SEQ ID NOS:8, 9 or 10. Alternatively, the invention may also be practiced with a polynucleotide consisting of the sequence of SEQ ID NOS:8, 9 or 10 in combination with one or more naturally occurring sequences that are normally found with SEQ ID NOS:8, 9 or 10.

Polynucleotides with sequences comprising SEQ ID NOS:8 or 9, either naturally occurring or synthetic, may be used to detect nucleic acids which are over expressed in breast cancer cells that are responsive to TAM treatment. Polynucleotides with sequences comprising SEQ ID NO:10, either naturally occurring or synthetic, may be used to detect nucleic acids which are under expressed in breast cancer cells that are responsive to TAM treatment.

Additional sequences that may be used in polynucleotides as described above for SEQ ID NOS:8 and 9 are the following:

SEQ ID NO:11:
TGCCTAATTTCACTCTCAGAGTGAGGCAGGTAACTGGGGCTCCACTGGGTCACTCTGAGA

SEQ ID NO:12:
TTGGAAGCAGAGTCCCTCTAAAGGTAACTCTTGTGGTCACTCAATATTGTATTGGCATTT

SEQ ID NO:13:
ACGTTAGACTTTTGCTGGCATTCAAGTCATGGCTAGTCTGTGTATTTAATAAATGTGTGT

SEQ ID NO:14:
CTGGTCAGCCACTCTGACTTTTCTACCACATTAAATTCTCCATTACATCTCACTATTGGT

SEQ ID NO:15:
TACAACTTCTGAATGCTGCACATTCTTCCAAAATGATCCTTAGCACAATCTATTGTATGA

SEQ ID NO:16:
GGGATGGCCTTTAGGCCACAGTAGTGTCTGTGTTAAGTTCACTAAATGTGTATTTAATGA

SEQ ID NO:17:
CTCAAAGTGCTAAAGCTATGGTTGACTGCTCTGGTGTTTTTATATTCATTCGTGCTTTAG

Additional sequences that may be used in polynucleotides as described above for SEQ ID NO:10 are the following:

SEQ ID NO:18:
CTATGGGGATGGTCCACTGTCACTGTTTCTCTGCTGTTGCAAATACATGGATAACACATT

SEQ ID NO:19:
ACTGGAAAAGCAGATGGTCTGACTGTGCTATGGCCTCATCATCAAGACTTTCAATCCTAT

SEQ ID NO:20:
ACGCCAAGCTCTTCAGTGAAGACACGATGTTATTAAAAGCCTGTTTTAGGGACTGCAAAA

SEQ ID NO:21:
TTTTTGTAAAATCTTTAACCTTCCCTTTGTTCTTCATGTACACGCTGAACTGCAATTCTT

SEQ ID NO:22:
AACCTGGGGCATTTAGGGCAGAGGACAAAAGGATGTCAGCAATTGCTTGGGCTGCTTGGC

SEQ ID NO:23:
CTGGAACCTCTGGACTCCCCATGCTCTAACTCCCACACTCTGCTATCAGAAACTTAAACT

SEQ ID NO:24:
AACCCCAGAACCATCTAAGACATGGGATTCAGTGATCATGTGGTTCTCCTTTTAACTTAC

SEQ ID NO:25:
GGCCATGTGCCATGGTATTTGGGTCCTGGGAGGGTGGGTGAAATAAAGGCATACTGTCTT

SEQ ID NO:26:
GTGTAGGCAGTCATGGCACCAAAGCCACCAGACTGACAAATGTGTATCAGATGCTTTTGT

SEQ ID NO:27:
GAAAACCTCTTCAAAAGACAAAAAGCTGGCACTGCATTCTCTCTCTGTAGCAGGACAGAA

SEQ ID NO:28:
CACATCTTTAGGGTCAGTGAACAATGGGGCACATTTGGCACTAGCTTGAGCCCAACTCTG

SEQ ID NO:29:
GCCTTAATTTCCTCATCTGAAAACTGGAAGGCCTGACTTGACTTGTTGAGCTTAAGATCC

SEQ ID NO:30:
CTTCAGGGGAGGATCAAGCTTTGAACCAAAGCCAATCACTGGCTTGATTTGTGTTTTTTA

SEQ ID NO:31:
ACAAGTTTTCACTGAATGAGCATGGCAGTGCCACTCAAGAAAATGAATCTCCAAAGTATC

Additionally, polynucleotides containing other sequences, particularly unique sequences, present in naturally occurring nucleic acid molecules comprising SEQ ID NOS:8-31 may be used in the practice of the invention.

Other polynucleotides for use in the practice of the invention include those that have sufficient homology to those described above to detect expression by use of hybridization techniques. Such polynucleotides preferably have about or 95%, about or 96%, about or 97%, about or 98%, or about or 99% identity with IL17BR, CACNA1D, or HOXB13 sequences as described herein. Identity is determined using the BLAST algorithm, as described above. The other polynucleotides for use in the practice of the invention may also be described on the basis of the ability to hybridize to polynucleotides of the invention under stringent conditions of about 30% v/v to about 50% formamide and from about 0.01M to about 0.15M salt for hybridization and from about 0.01M to about 0.15M salt for wash conditions at about 55 to about 65°C or higher, or conditions equivalent thereto.

In a further embodiment of the invention, a population of single stranded nucleic acid molecules comprising one or both strands of a human IL17BR or CACNA1D sequence is provided as a probe such that at least a portion of said population may be hybridized to one or both strands of a nucleic acid molecule quantitatively amplified from RNA of a breast cancer cell. The population may be only the antisense strand of a human IL17BR or CACNA1D sequence such that a sense strand of a molecule from, or amplified from, a breast cancer cell may be hybridized to a portion of said population. The population preferably comprises a sufficiently excess amount of said one or both strands of a human IL17BR or CACNA1D sequence in comparison to the amount of expressed (or amplified) nucleic acid molecules containing a complementary IL17BR or CACNA1D sequence from a normal breast cell. This condition of excess permits the increased amount of nucleic acid expression in a breast cancer cell to be readily detectable as an increase.

Alternatively, the population of single stranded molecules is equal to or in excess of all of one or both strands of the nucleic acid molecules amplified from a breast cancer cell such that the population is sufficient to hybridize to all of one or both strands. Preferred cells are those of a breast cancer patient that is ER+ or for whom tamoxifen treatment is contemplated. The single stranded molecules may of course be the denatured form of any IL17BR and/or CACNA1D sequence containing double stranded nucleic acid molecule or polynucleotide as described herein.

The population may also be described as being hybridized to IL17BR or CACNA1D sequence containing nucleic acid molecules at a level of at least twice as much as that by nucleic

acid molecules of a normal breast cell. As in the embodiments described above, the nucleic acid molecules may be those quantitatively amplified from a breast cancer cell such that they reflect the amount of expression in said cell.

The population is preferably immobilized on a solid support, optionally in the form of a location on a microarray. A portion of the population is preferably hybridized to nucleic acid molecules quantitatively amplified from a non-normal or abnormal breast cell by real time PCR. The real time PCR may be practiced by use of amplified RNA from a breast cancer cell, as long as the amplification used was quantitative with respect to IL17BR or CACNA1D containing sequences.

In another embodiment of the invention, expression based on detection of DNA status may be used. Detection of the HOXB13 DNA as methylated, deleted or otherwise inactivated, may be used as an indication of decreased expression as found in non-normal breast cells. This may be readily performed by PCR based methods known in the art. The status of the promoter regions of HOXB13 may also be assayed as an indication of decreased expression of HOXB13 sequences. A non-limiting example is the methylation status of sequences found in the promoter region.

Conversely, detection of the DNA of a sequence as amplified may be used for as an indication of increased expression as found in non-normal breast cells. This may be readily performed by PCR based, fluorescent *in situ* hybridization (FISH) and chromosome *in situ* hybridization (CISH) methods known in the art.

A preferred embodiment using a nucleic acid based assay to determine expression is by immobilization of one or more of the sequences identified herein on a solid support, including, but not limited to, a solid substrate as an array or to beads or bead based technology as known in the art. Alternatively, solution based expression assays known in the art may also be used. The immobilized sequence(s) may be in the form of polynucleotides as described herein such that the polynucleotide would be capable of hybridizing to a DNA or RNA corresponding to the sequence(s).

The immobilized polynucleotide(s) may be used to determine the state of nucleic acid samples prepared from sample breast cancer cell(s), optionally as part of a method to detect ER status in said cell(s). Without limiting the invention, such a cell may be from a patient suspected of

being afflicted with, or at risk of developing, breast cancer. The immobilized polynucleotide(s) need only be sufficient to specifically hybridize to the corresponding nucleic acid molecules derived from the sample (and to the exclusion of detectable or significant hybridization to other nucleic acid molecules).

In yet another embodiment of the invention, a ratio of the expression levels of two of the disclosed genes may be used to predict response to TAM treatment. Preferably, the ratio is that of two genes with opposing patterns of expression, such as an underexpressed gene to an overexpressed gene. Non-limiting examples include the ratio of HOXB13 over IL17BR or the ratio of HOXB13 over CACNA1D. This aspect of the invention is based in part on the observation that such a ratio has a stronger correlation with TAM treatment outcome than the expression level of either gene alone. For example, the ratio of HOXB13 over IL17BR has an observed classification accuracy of 77%.

Additional Embodiments of the Invention

In embodiments where only one or a few genes are to be analyzed, the nucleic acid derived from the sample breast cancer cell(s) may be preferentially amplified by use of appropriate primers such that only the genes to be analyzed are amplified to reduce contaminating background signals from other genes expressed in the breast cell. Alternatively, and where multiple genes are to be analyzed or where very few cells (or one cell) is used, the nucleic acid from the sample may be globally amplified before hybridization to the immobilized polynucleotides. Of course RNA, or the cDNA counterpart thereof may be directly labeled and used, without amplification, by methods known in the art.

Sequence expression based on detection of a presence, increase, or decrease in protein levels or activity may also be used. Detection may be performed by any immunohistochemistry (IHC) based, bodily fluid based (where a IL17BR, CACNA1D, and/or HOXB13 polypeptide is found in a bodily fluid, such as but not limited to blood), antibody (including autoantibodies against the protein where present) based, ex foliate cell (from the cancer) based, mass spectroscopy based, and image (including used of labeled ligand where available) based method known in the art and recognized as appropriate for the detection of the protein. Antibody and image based methods are

additionally useful for the localization of tumors after determination of cancer by use of cells obtained by a non-invasive procedure (such as ductal lavage or fine needle aspiration), where the source of the cancerous cells is not known. A labeled antibody or ligand may be used to localize the carcinoma(s) within a patient.

Antibodies for use in such methods of detection include polyclonal antibodies, optionally isolated from naturally occurring sources where available, and monoclonal antibodies, including those prepared by use of IL17BR, CACNA1D, and/or HOXB13 polypeptides as antigens. Such antibodies, as well as fragments thereof (including but not limited to F_{ab} fragments) function to detect or diagnose non-normal or cancerous breast cells by virtue of their ability to specifically bind IL17BR, CACNA1D, or HOXB13 polypeptides to the exclusion of other polypeptides to produce a detectable signal. Recombinant, synthetic, and hybrid antibodies with the same ability may also be used in the practice of the invention. Antibodies may be readily generated by immunization with a IL17BR, CACNA1D, or HOXB13 polypeptide, and polyclonal sera may also be used in the practice of the invention.

Antibody based detection methods are well known in the art and include sandwich and ELISA assays as well as Western blot and flow cytometry based assays as non-limiting examples. Samples for analysis in such methods include any that contain IL17BR, CACNA1D, or HOXB13 polypeptides. Non-limiting examples include those containing breast cells and cell contents as well as bodily fluids (including blood, serum, saliva, lymphatic fluid, as well as mucosal and other cellular secretions as non-limiting examples) that contain the polypeptides.

The above assay embodiments may be used in a number of different ways to identify or detect the response to TAM treatment based on gene expression in a breast cancer cell sample from a patient. In some cases, this would reflect a secondary screen for the patient, who may have already undergone mammography or physical exam as a primary screen. If positive from the primary screen, the subsequent needle biopsy, ductal lavage, fine needle aspiration, or other analogous methods may provide the sample for use in the assay embodiments before, simultaneous with, or after assaying for ER status. The present invention is particularly useful in combination with non-invasive protocols, such as ductal lavage or fine needle aspiration, to prepare a breast cell sample.

The present invention provides a more objective set of criteria, in the form of gene expression profiles of a discrete set of genes, to discriminate (or delineate) between breast cancer outcomes. In particularly preferred embodiments of the invention, the assays are used to discriminate between good and poor outcomes after tamoxifen treatment. Comparisons that discriminate between outcomes after about 10, about 20, about 30, about 40, about 50, about 60, about 70, about 80, about 90, about 100, or about 150 months may be performed.

While good and poor survival outcomes may be defined relatively in comparison to each other, a "good" outcome may be viewed as a better than 50% survival rate after about 60 months post surgical intervention to remove breast cancer tumor(s). A "good" outcome may also be a better than about 60%, about 70%, about 80% or about 90% survival rate after about 60 months post surgical intervention. A "poor" outcome may be viewed as a 50% or less survival rate after about 60 months post surgical intervention to remove breast cancer tumor(s). A "poor" outcome may also be about a 70% or less survival rate after about 40 months, or about a 80% or less survival rate after about 20 months, post surgical intervention.

In another embodiment of the invention based on the expression of multiple genes in an expression pattern or profile, the isolation and analysis of a breast cancer cell sample may be performed as follows:

(1) Ductal lavage or other non-invasive procedure is performed on a patient to obtain a sample.

(2) Sample is prepared and coated onto a microscope slide. Note that ductal lavage results in clusters of cells that are cytologically examined as stated above.

(3) Pathologist or image analysis software scans the sample for the presence of non-normal and/or atypical breast cancer cells.

(4) If such cells are observed, those cells are harvested (e.g. by microdissection such as LCM).

(5) RNA is extracted from the harvested cells.

(6) RNA is purified, amplified, and labeled.

(7) Labeled nucleic acid is contacted with a microarray containing polynucleotides of the genes identified herein as correlated to discriminations between breast cancer outcomes under suitable hybridization conditions, then processed and scanned to obtain a pattern of

intensities of each spot (relative to a control for general gene expression in cells) which determine the level of expression of the gene(s) in the cells.

(8) The pattern of intensities is analyzed by comparison to the expression patterns of the genes in known samples of breast cancer cells correlated with outcomes (relative to the same control).

A specific example of the above method would be performing ductal lavage following a primary screen, observing and collecting non-normal and/or atypical cells for analysis. The comparison to known expression patterns, such as that made possible by a model generated by an algorithm (such as, but not limited to nearest neighbor type analysis, SVM, or neural networks) with reference gene expression data for the different breast cancer survival outcomes, identifies the cells as being correlated with subjects with good or poor outcomes. Another example would be taking a breast tumor removed from a subject after surgical intervention, optionally converting all or part of it to an FFPE sample prior to subsequent isolation and preparation of breast cancer cells from the tumor for determination/identification of atypical, non-normal, or cancer cells, and isolation of said cells followed by steps 5 through 8 above.

Alternatively, the sample may permit the collection of both normal as well as cancer cells for analysis. The gene expression patterns for each of these two samples will be compared to each other as well as the model and the normal versus individual comparisons therein based upon the reference data set. This approach can be significantly more powerful that the cancer cells only approach because it utilizes significantly more information from the normal cells and the differences between normal and cancer cells (in both the sample and reference data sets) to determine the breast cancer outcome of the patient based on gene expression in the cancer cells from the sample.

In yet another embodiment of the invention based on the expression of a few genes, the isolation and analysis of a breast cancer cell sample may be performed as follows:

(1) Ductal lavage or other non-invasive procedure is performed on a patient to obtain a sample.

(2) Sample is prepared and coated onto a microscope slide. Note that ductal lavage results in clusters of cells that are cytologically examined as stated above.

(3) Pathologist or image analysis software scans the sample for the presence of atypical cells.

(4) If atypical cells are observed, those cells are harvested (e.g. by microdissection such as LCM).

(5) RNA is extracted from the harvested cells.

(6) RNA is assayed, directly or after conversion to cDNA or amplification therefrom, for the expression of IL17BR, CACNA1D, and/or HOXB13 sequences.

One example of the above method would be performing ductal lavage following a primary screen, observing and collecting non-normal cells (or cells suspected of being non-normal) for analysis. Alternatively, the sample may permit the collection of both normal and non-normal cells (or cells suspected of being non-normal) for analysis. The expression levels of IL17BR, CACNA1D, and/or HOXB13 sequences in each of these two populations may be compared to each other. This approach can be significantly more powerful than one using the non-normal cells only approach because it utilizes information from the normal cells and the differences between normal and non-normal cells to determine the status of the non-normal cells from the sample.

With use of the present invention, skilled physicians may prescribe or withhold TAM treatment based on prognosis determined via practice of the instant invention.

The above discussion is also applicable where a palpable lesion is detected followed by fine needle aspiration or needle biopsy of cells from the breast. The cells are plated and reviewed by a pathologist or automated imaging system which selects cells for analysis as described above.

The present invention may also be used, however, with solid tissue biopsies, including those stored as an FFPE specimen. For example, a solid biopsy may be collected and prepared for visualization followed by determination of expression of one or more genes identified herein to determine the breast cancer outcome. As another non-limiting example, a solid biopsy may be collected and prepared for visualization followed by determination of increased IL17BR and/or CACNA1D expression. One preferred means is by use of *in situ* hybridization with polynucleotide or protein identifying probe(s) for assaying expression of said gene(s). An analogous method may be used to detect decreased expression of HOXB13 sequences.

In an alternative method, the solid tissue biopsy may be used to extract molecules followed by analysis for expression of one or more gene(s). This provides the possibility of leaving out the

need for visualization and collection of only cancer cells or cells suspected of being cancerous. This method may of course be modified such that only cells that have been positively selected are collected and used to extract molecules for analysis. This would require visualization and selection as a prerequisite to gene expression analysis. In the case of an FFPE sample, cells may be obtained followed by RNA extraction, amplification and detection as described herein.

In a further modification of the above, both normal cells and cancer cells are collected and used to extract molecules for analysis of gene expression. The approach, benefits and results are as described above using non-invasive sampling.

In a further alternative to all of the above, the sequence(s) identified herein may be used as part of a simple PCR or array based assay simply to determine the response to TAM treatment by use of a sample from a non-invasive sampling procedure. The detection of sequence expression from samples may be by use of a single microarray able to assay expression of the disclosed sequences as well as other sequences, including sequences known not to vary in expression levels between normal and non-normal breast cells, for convenience and improved accuracy.

Other uses of the present invention include providing the ability to identify breast cancer cell samples as having different responses to TAM treatment for further research or study. This provides an advance based on objective genetic/molecular criteria.

The genes identified herein also may be used to generate a model capable of predicting the breast cancer survival and recurrence outcomes of an ER+ breast cell sample based on the expression of the identified genes in the sample. Such a model may be generated by any of the algorithms described herein or otherwise known in the art as well as those recognized as equivalent in the art using gene(s) (and subsets thereof) disclosed herein for the identification of breast cancer outcomes. The model provides a means for comparing expression profiles of gene(s) of the subset from the sample against the profiles of reference data used to build the model. The model can compare the sample profile against each of the reference profiles or against a model defining delineations made based upon the reference profiles. Additionally, relative values from the sample profile may be used in comparison with the model or reference profiles.

In a preferred embodiment of the invention, breast cell samples identified as normal and cancerous from the same subject may be analyzed, optionally by use of a single microarray, for

their expression profiles of the genes used to generate the model. This provides an advantageous means of identifying survival and recurrence outcomes based on relative differences from the expression profile of the normal sample. These differences can then be used in comparison to differences between normal and individual cancerous reference data which was also used to generate the model.

Articles of Manufacture

The materials and methods of the present invention are ideally suited for preparation of kits produced in accordance with well known procedures. The invention thus provides kits comprising agents (like the polynucleotides and/or antibodies described herein as non-limiting examples) for the detection of expression of the disclosed sequences. Such kits, optionally comprising the agent with an identifying description or label or instructions relating to their use in the methods of the present invention, are provided. Such a kit may comprise containers, each with one or more of the various reagents (typically in concentrated form) utilized in the methods, including, for example, pre-fabricated microarrays, buffers, the appropriate nucleotide triphosphates (e.g., dATP, dCTP, dGTP and dTTP; or rATP, rCTP, rGTP and UTP), reverse transcriptase, DNA polymerase, RNA polymerase, and one or more primer complexes of the present invention (e.g., appropriate length poly(T) or random primers linked to a promoter reactive with the RNA polymerase). A set of instructions will also typically be included.

The methods provided by the present invention may also be automated in whole or in part. All aspects of the present invention may also be practiced such that they consist essentially of a subset of the disclosed genes to the exclusion of material irrelevant to the identification of breast cancer survival outcomes via a cell containing sample.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

Examples

Example 1

Gene expression signature predicting TAM treatment outcome in breast cancer

A cohort of 62 estrogen receptor-positive breast cancer patients were uniformly treated with the anti-estrogen drug tamoxifen (TAM), and followed for up to 14 years. 33 patients recurred whereas 29 patients remained disease-free during the entire follow up periods. Correlating gene expression patterns with tumor recurrence/non-recurrence, a set of genes was discovered whose expression levels differ significantly between these two groups. This gene expression signature can thus be used to predict whether a patient will respond to TAM as first-line treatment based on the gene expression profile of a routine biopsy of the primary cancer.

Laser capture microdissection was performed on each tumor biopsy to procure pure populations of cancerous epithelial cells, which were then analyzed on a 22000-gene high-density oligonucleotide microarray. The top 25% genes with the greatest variances across all samples (n=5475) were selected for signature extraction. Genes showing statistically significant correlations with tumor recurrence/non-recurrence were identified using two different statistical techniques.

In the first approach, patients were divided into two groups (recurrence vs. non-recurrence), and a standard t-test was performed for each gene, which identified 149 genes with p values < 0.001. The results for this analysis are shown in Table 1. Genes identified by their accession numbers correlate with non-responders when the t-statistic is less than zero while genes with a t-statistic greater than zero correlate to positive responders.

Table 1. 149-gene signature identified by t-test

Accession	p value	t-statistic	Description
BC002595	5.49E-10	-8.186189	NDUFB7 | NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 7 (18kD, B18)
BC002705	1.65E-09	-7.550191	C22orf3 | chromosome 22 open reading frame 3
AL080126	1.82E-09	-7.410723	KIAA0683 | KIAA0683 gene product
AI767799	2.02E-09	-7.768777	BBC3 | Bcl-2 binding component 3
AL021683	2.78E-09	-7.083131	SCO2 | SCO cytochrome oxidase deficient homolog 2 (yeast)
BC000507	4.38E-09	-7.026423	MAAT1 | melanoma-associated antigen recognised by cytotoxic T

Accession	p value	t-statistic	Description
			lymphocytes
AK027124	1.70E-08	-6.740214	FLJ23471 | hypothetical protein FLJ23471
BC016737	1.99E-08	-6.742271	MPST | mercaptopyruvate sulfurtransferase
BC011874	3.53E-08	-6.327036	MGC20486 | hypothetical protein MGC20486
			HMGIY | high-mobility group (nonhistone chromosomal) protein isoforms I
BC008832	3.86E-08	-6.388736	and Y
			NDUFS6 | NADH dehydrogenase (ubiquinone) Fe-S protein 6 (13kD)
AF044959	5.20E-08	-6.222993	(NADH-coenzyme Q reductase)
BC016832	6.61E-08	-6.627917	MGC4607 | hypothetical protein MGC4607
BC011680	6.61E-08	-6.427017	DKFZp434G0522 | hypothetical protein DKFZp434G0522
AA811922	6.75E-08	-6.634444	FLJ10140 | hypothetical protein FLJ10140
AW075691	1.03E-07	-6.272638	KIAA1847 | hypothetical protein FLJ14972
AK024627	1.14E-07	-6.019024	FLJ20974 | hypothetical protein FLJ20974
			ATP5D | ATP synthase, H+ transporting, mitochondrial F1 complex, delta
BC002389	1.15E-07	-6.05372	subunit
			Homo sapiens cDNA FLJ30733 fis, clone FEBRA2000129, moderately similar
			to PROBABLE TRNA (5-METHYLAMINOMETHYL-2-
AK055295	1.24E-07	-6.391213	THIOURIDYLATE)-METH YLTRANSFERASE (EC 2.1.1.61)
BC011621	1.54E-07	5.943998	HOOK1 | hook1 protein
AK023601	1.69E-07	5.919878	Homo sapiens cDNA FLJ13539 fis, clone PLACE1006640
BC013959	1.83E-07	-6.09348	GNL1 | guanine nucleotide binding protein-like 1
BC018346	1.84E-07	-5.929725	LAK-4P | expressed in activated T/LAK lymphocytes
AF052052	3.46E-07	-5.920813	TFPT | TCF3 (E2A) fusion partner (in childhood Leukemia)
AL136921	3.66E-07	-5.742098	DKFZp586I021 | hypothetical protein DKFZp586I021
AI968598	6.33E-07	-5.685799	Homo sapiens cDNA FLJ12182 fis, clone MAMMA1000761
			ERP70 | protein disulfide isomerase related protein (calcium-binding protein,
BC011754	7.93E-07	-5.671882	intestinal-related)
BC014270	3.58E-06	-5.155079	PRKCZ | protein kinase C, zeta
NM_001130	3.82E-06	-5.120513	AES | amino-terminal enhancer of split
BF116098	4.09E-06	5.101295	ESTs
BC015594	5.01E-06	-5.027872	Homo sapiens mRNA for FLJ00083 protein, partial cds
AK000081	5.74E-06	-4.996636	CDC2L1 | cell division cycle 2-like 1 (PITSLRE proteins)
NM_006278	6.23E-06	-4.968186	SIAT4C | sialyltransferase 4C (beta-galactosidase alpha-2,3-sialytransferase)
BC008841	6.32E-06	-5.039493	KIAA0415 | KIAA0415 gene product
			Homo sapiens cDNA FLJ32384 fis, clone SKMUS1000104, weakly similar to
AI972367	7.05E-06	-4.93464	Homo sapiens mRNA for HEXIM1 protein, complete cds
AI467849	7.34E-06	-4.933176	TBC1D1 | TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1
			QPRT | quinolinate phosphoribosyltransferase (nicotinate-nucleotide
NM_014298	9.19E-06	-4.869139	pyrophosphorylase (carboxylating))
			ESTs, Weakly similar to JC5238 galactosylceramide-like protein, GCP
H19223	1.15E-05	4.786877	[H.sapiens]
AI638324	1.22E-05	4.783615	Homo sapiens cDNA FLJ30332 fis, clone BRACE2007254
AF208111	1.30E-05	4.761353	IL17BR | interleukin 17B receptor
NM_020978	1.34E-05	4.803041	AMY2B | amylase, alpha 2B; pancreatic
BC015497	1.59E-05	-4.722392	TEAD4 | TEA domain family member 4
AI561249	1.69E-05	4.681189	KTN1 | kinectin 1 (kinesin receptor)

Accession	p value	t-statistic	Descripti n
BC004235	1.73E-05	-4.684545	DDX38 | DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 38
NM_013347	1.89E-05	4.67568	HSU24186 | replication protein A complex 34 kd subunit homolog Rpa4
AL117616	1.90E-05	4.645713	SRI | sorcin
AL117478	2.00E-05	-4.634086	AGS3 | likely ortholog of rat activator of G-protein signaling 3
NM_006304	2.28E-05	4.59794	DSS1 | Deleted in split-hand/split-foot 1 region
BC009507	2.29E-05	-4.59323	ISG15 | interferon-stimulated protein, 15 kDa
AK025141	2.89E-05	4.529022	Homo sapiens cDNA: FLJ21488 fis, clone COL05445
AA581602	4.04E-05	4.43179	ESTs
BC006499	4.22E-05	-4.422009	HRAS | v-Ha-ras Harvey rat sarcoma viral oncogene homolog
BC007066	5.23E-05	4.379391	CDA11 | CDA11 protein
BC009869	5.35E-05	4.352129	SERF2 | small EDRK-rich factor 2
AA206609	5.68E-05	-4.339494	Homo sapiens cDNA FLJ30002 fis, clone 3NB691000085
AI682928	5.76E-05	4.350598	EST
BC006284	7.29E-05	-4.359234	Homo sapiens, clone IMAGE:3957135, mRNA, partial cds
AI871458	7.41E-05	-4.303954	ESTs
AF068918	7.50E-05	-4.284961	BIN1 | bridging integrator 1
NM_018936	7.50E-05	-4.254075	PCDHB2 | protocadherin beta 2
AI469557	7.83E-05	-4.248879	EPHB3 | EphB3
AL137521	8.02E-05	-4.27827	Homo sapiens mRNA; cDNA DKFZp434D0218 (from clone DKFZp434D0218); partial cds
AI268007	8.04E-05	4.245279	Homo sapiens cDNA FLJ30137 fis, clone BRACE2000078
AW070918	8.56E-05	-4.21829	ESTs, Weakly similar to T2D3_HUMAN TRANSCRIPTION INITIATION FACTOR TFIID 135 KDA SUBUNIT [H.sapiens]
AK025862	8.75E-05	4.237223	Homo sapiens cDNA: FLJ22209 fis, clone HRC01496
AI264644	9.54E-05	-4.240955	KIAA0775 | KIAA0775 gene product
BF438928	9.75E-05	4.180144	ESTs
BC001403	9.83E-05	-4.17366	CPSF5 | cleavage and polyadenylation specific factor 5, 25 kD subunit
AI270018	1.01E-04	-4.167464	ECE1 | endothelin converting enzyme 1
AL133427	1.04E-04	4.19331	Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 261172
AI400775	1.12E-04	-4.148062	RABL2B | RAB, member of RAS oncogene family-like 2B
AW016075	1.21E-04	4.132864	ESTs, Weakly similar to ALUA_HUMAN !!!! ALU CLASS A WARNING ENTRY !!! [H.sapiens]
AI033912	1.26E-04	4.100849	RLN2 | relaxin 2 (H2)
AA668884	1.28E-04	4.104243	ESTs
AL133661	1.38E-04	4.085685	DKFZp434C0328 | hypothetical protein DKFZp434C0328
BC009874	1.40E-04	-4.074407	JUN | v-jun sarcoma virus 17 oncogene homolog (avian)
AI357434	1.52E-04	4.055067	HSP105B | heat shock 105kD
AF119871	1.54E-04	4.081889	PRO2268 | hypothetical protein PRO2268
AK024715	1.54E-04	4.043172	FLJ21062 | hypothetical protein FLJ21062
X62534	1.58E-04	4.048006	HMG2 | high-mobility group (nonhistone chromosomal) protein 2
BI793002	1.60E-04	4.039819	OSBPL8 | oxysterol binding protein-like 8
L13738	1.61E-04	-4.041465	ACK1 | activated p21cdc42Hs kinase
AW297123	1.74E-04	4.019412	ESTs
NM_020235	1.80E-04	4.011596	BBX | bobby sox homolog (Drosophila)

Accession	p value	t-statistic	Description
AI686003	1.83E-04	4.035297	ESTs
AK022916	1.84E-04	3.989755	ZNF281 | zinc finger protein 281
AK025701	1.86E-04	-3.99009	PLXNB2 | plexin B2
AA806831	1.91E-04	-4.126686	ESTs
AL117396	1.93E-04	3.982093	DKFZP586M0622 | DKFZP586M0622 protein
AW192535	1.93E-04	3.982278	ESTs
AW076080	1.94E-04	3.972626	Homo sapiens, clone IMAGE:3463399, mRNA, partial cds
AB014541	1.95E-04	-3.97255	AATK | apoptosis-associated tyrosine kinase
AK024967	1.96E-04	4.008564	Homo sapiens cDNA: FLJ21314 fis, clone COL02248
BC018644	2.10E-04	-3.981862	NUDT8 | nudix (nucleoside diphosphate linked moiety X)-type motif 8
AK026817	2.11E-04	3.9468	FLJ23577 | hypothetical protein FLJ23577
BC000692	2.20E-04	-3.943535	HYAL2 | hyaluronoglucosaminidase 2
BE967259	2.26E-04	3.927279	BCL2 | B-cell CLL/lymphoma 2
NM_004038	2.29E-04	3.946754	AMY1A | amylase, alpha 1A; salivary
AF052110	2.34E-04	-3.915428	DAF | decay accelerating factor for complement (CD55, Cromer blood group system)
AW069725	2.38E-04	3.914238	CRYZ | crystallin, zeta (quinone reductase)
BM127867	2.44E-04	3.908237	MDM1 | nuclear protein double minute 1
AL050227	2.50E-04	3.894782	Homo sapiens mRNA; cDNA DKFZp586M0723 (from clone DKFZp586M0723)
BC005377	2.61E-04	3.949255	ACADM | acyl-Coenzyme A dehydrogenase, C-4 to C-12 straight chain
BC006437	2.66E-04	-3.880036	C321D2.4 | hypothetical protein C321D2.4
AF153330	2.73E-04	3.871579	SLC19A2 | solute carrier family 19 (thiamine transporter), member 2
AA635853	2.86E-04	3.856068	EST
AK021798	2.92E-04	3.858723	FLJ11736 | hypothetical protein FLJ11736
BE675157	3.06E-04	3.882041	ESTs
T52873	3.08E-04	3.831368	ESTs, Moderately similar to G02075 transcription repressor zinc finger protein 85 [H.sapiens]
BE645958	3.30E-04	3.812843	ESTs
BF589163	3.37E-04	3.857405	ESTs
AA040945	3.44E-04	-3.797113	ESTs
AK001783	3.74E-04	3.771144	FLJ10921 | hypothetical protein FLJ10921
R43003	4.06E-04	3.80021	ESTs, Highly similar to COBW-like protein [H.sapiens]
AW135596	4.10E-04	3.742774	FLJ10058 | hypothetical protein FLJ10058
NM_003489	4.20E-04	3.736095	NRIP1 | nuclear receptor interacting protein 1
AL136663	4.25E-04	-3.748587	DKFZp564A176 | hypothetical protein DKFZp564A176
AI376433	4.47E-04	3.774197	KIAA1912 | KIAA1912 protein
BC015792	4.49E-04	-3.725478	Homo sapiens, clone MGC:23665 IMAGE:4866941, mRNA, complete cds
AI478784	4.63E-04	3.705085	FLJ11267 | hypothetical protein FLJ11267
U50532	4.91E-04	3.723884	CG005 | hypothetical protein from BCRA2 region
AI700363	4.92E-04	-3.719752	ESTs
BC005956	5.22E-04	3.679274	RLN1 | relaxin 1 (H1)
AI240933	5.44E-04	3.657963	ESTs
AF330046	5.51E-04	3.652748	PIBF1 | progesterone-induced blocking factor 1

Accession	p value	t-statistic	Description
AI128331	5.55E-04	3.648721	ENDOFIN | endosome-associated FYVE-domain protein
BC008381	5.63E-04	3.654514	IMPA1 | inositol(myo)-1(or 4)-monophosphatase 1
AF023676	5.64E-04	-3.647402	TM7SF2 | transmembrane 7 superfamily member 2
AL050179	5.73E-04	3.665736	TPM1 | tropomyosin 1 (alpha)
BC002355	5.73E-04	3.654105	HNRPA1 | heterogeneous nuclear ribonucleoprotein A1
AK056075	5.84E-04	3.632268	Homo sapiens cDNA FLJ31513 fis, clone NT2RI1000127
AK024999	6.01E-04	3.641434	Homo sapiens cDNA: FLJ21346 fis, clone COL02705
AK000305	6.30E-04	3.666154	FLJ20298 | hypothetical protein FLJ20298
AF085243	6.47E-04	3.601667	ZNF236 | zinc finger protein 236
AW510501	6.56E-04	3.620023	ARHGAP5 | Rho GTPase activating protein 5
AI953054	6.57E-04	-3.59919	TKT | transketolase (Wernicke-Korsakoff syndrome)
BC012628	7.09E-04	-3.610827	TCAP | titin-cap (telethonin)
BC007092	7.12E-04	-3.598786	HOXB13 | homeo box B13
AB000520	7.40E-04	-3.558109	APS | adaptor protein with pleckstrin homology and src homology 2 domains
AW150267	7.47E-04	3.566503	C21orf9 | chromosome 21 open reading frame 9
AI800042	7.64E-04	3.575129	ESTs
AF033199	8.01E-04	-3.541312	ZNF204 | zinc finger protein 204
BC002607	8.15E-04	-3.529271	KIAA1446 | KIAA1446 protein
BC002480	8.43E-04	-3.525938	FLJ13352 | hypothetical protein FLJ13352
AI568728	9.04E-04	-3.501174	SKI | v-ski sarcoma viral oncogene homolog (avian)
AA648536	9.20E-04	-3.48714	MYO1E | myosin IE
AI335002	9.28E-04	3.502278	PBEF | pre-B-cell colony-enhancing factor
AW452172	9.45E-04	3.483191	ESTs
AF334676	9.50E-04	3.476947	TEKT3 | tektin 3
AF085233	9.77E-04	3.479809	SGKL | serum/glucocorticoid regulated kinase-like

In the second approach, the actual times of recurrence or follow-up (for those who remained disease-free) were used in a Cox proportional hazard regression model using each gene as the single predictor variable, identifying 149 genes with p values (Wald statistic) < 0.001. The results for this analysis are shown in Table 2. Genes identified by their accession numbers correlate with subjects likely to suffer a reoccurrence after TAM therapy when the hazard ratio is greater than one while genes with a hazard ration of less than one correlate to individuals who are likely not to suffer a reoccurrence of breast cancer.

Table 2. 149-gene signature identified by Cox regression

Accession	p value	hazard ratio	Description
BC002595	3.00E-08	1.9899702	NDUFB7 | NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 7 (18kD, B18)
BC000507	3.66E-08	2.3494974	MAAT1 | melanoma-associated antigen recognised by cytotoxic T lymphocytes
BC016832	5.45E-08	2.2890356	MGC4607 | hypothetical protein MGC4607
BC002705	1.52E-07	2.5669791	C22orf3 | chromosome 22 open reading frame 3
AI767799	1.93E-07	2.1989649	BBC3 | Bcl-2 binding component 3
BC011874	2.51E-07	2.8556338	MGC20486 | hypothetical protein MGC20486
AL021683	3.74E-07	2.1946935	SCO2 | SCO cytochrome oxidase deficient homolog 2 (yeast)
BC008832	4.28E-07	2.3960849	HMGIY | high-mobility group (nonhistone chromosomal) protein isoforms I and Y
AL080126	4.46E-07	2.1613379	KIAA0683 | KIAA0683 gene product
BC013959	4.68E-07	2.4974081	GNL1 | guanine nucleotide binding protein-like 1
AF052052	5.29E-07	2.1949663	TFPT | TCF3 (E2A) fusion partner (in childhood Leukemia)
AA811922	6.00E-07	1.9841656	FLJ10140 | hypothetical protein FLJ10140
BC011680	6.96E-07	2.373463	DKFZp434G0522 | hypothetical protein DKFZp434G0522
BC016737	1.06E-06	1.8482073	MPST | mercaptopyruvate sulfurtransferase
AI968598	1.24E-06	2.6284635	Homo sapiens cDNA FLJ12182 fis, clone MAMMA1000761
AW075691	1.35E-06	2.0681292	KIAA1847 | hypothetical protein FLJ14972
AK024627	1.53E-06	2.6015319	FLJ20974 | hypothetical protein FLJ20974
AF044959	1.56E-06	2.8966077	NDUFS6 | NADH dehydrogenase (ubiquinone) Fe-S protein 6 (13kD) (NADH-coenzyme Q reductase)
BC002389	1.64E-06	1.8888501	ATP5D | ATP synthase, H+ transporting, mitochondrial F1 complex, delta subunit
AK055295	3.03E-06	1.8815611	Homo sapiens cDNA FLJ30733 fis, clone FEBRA2000129, moderately similar to PROBABLE TRNA (5-METHYLAMINOMETHYL-2-THIOURIDYLATE)-METH YLTRANSFERASE (EC 2.1.1.61)
BC005377	3.41E-06	0.5676057	ACADM | acyl-Coenzyme A dehydrogenase, C-4 to C-12 straight chain
H19223	4.47E-06	0.4802045	ESTs, Weakly similar to JC5238 galactosylceramide-like protein, GCP [H.sapiens]
AK023601	4.81E-06	0.4390305	Homo sapiens cDNA FLJ13539 fis, clone PLACE1006640
NM_001130	5.72E-06	2.1351138	AES | amino-terminal enhancer of split
NM_014298	6.39E-06	1.8007172	QPRT | quinolinate phosphoribosyltransferase (nicotinate-nucleotide pyrophosphorylase (carboxylating))
AK027124	7.12E-06	1.968632	FLJ23471 | hypothetical protein FLJ23471
AL117396	7.58E-06	0.4156321	DKFZP586M0622 | DKFZP586M0622 protein
AL136921	8.27E-06	2.3643799	DKFZp586I021 | hypothetical protein DKFZp586I021
U50532	8.81E-06	0.4216183	CG005 | hypothetical protein from BCRA2 region
BC018346	1.14E-05	1.8491373	LAK-4P | expressed in activated T/LAK lymphocytes
NM_013347	1.35E-05	0.3648298	HSU24186 | replication protein A complex 34 kd subunit

Accession	p value	hazard ratio	Description
			homolog Rpa4
BC011621	1.37E-05	0.5264059	HOOK1 | hook1 protein
BC006284	1.48E-05	2.1550372	Homo sapiens, clone IMAGE:3957135, mRNA, partial cds
BC004235	2.01E-05	2.4910338	DDX38 | DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 38
NM_006278	2.06E-05	1.9872895	SIAT4C | sialyltransferase 4C (beta-galactosidase alpha-2,3-sialytransferase)
AI972367	2.13E-05	2.1500078	Homo sapiens cDNA FLJ32384 fis, clone SKMUS1000104, weakly similar to Homo sapiens mRNA for HEXIM1 protein, complete cds
BC012628	2.31E-05	2.0388066	TCAP | titin-cap (telethonin)
AA581602	2.44E-05	0.4839842	ESTs
NM_018936	2.46E-05	1.4853858	PCDHB2 | protocadherin beta 2
AA746504	2.68E-05	0.667095	Homo sapiens cDNA FLJ30188 fis, clone BRACE2001267
AF220030	2.73E-05	0.4441676	TRIM6 | tripartite motif-containing 6
AI682928	2.90E-05	0.4144403	EST
AA206609	3.05E-05	2.0738914	Homo sapiens cDNA FLJ30002 fis, clone 3NB691000085
AL117616	3.06E-05	0.5506486	SRI | sorcin
U08997	3.06E-05	0.548039	GLUD2 | Glutamate dehydrogenase-2
BC009869	3.17E-05	0.4884412	SERF2 | small EDRK-rich factor 2
AL137521	3.24E-05	2.4199381	Homo sapiens mRNA; cDNA DKFZp434D0218 (from clone DKFZp434D0218); partial cds
AI871458	3.26E-05	2.0738428	ESTs
BC008841	3.27E-05	1.8195551	KIAA0415 | KIAA0415 gene product
AI467849	4.07E-05	1.689976	TBC1D1 | TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1
BC011754	4.42E-05	1.6224459	ERP70 | protein disulfide isomerase related protein (calcium-binding protein, intestinal-related)
AL050227	4.44E-05	0.7135796	Homo sapiens mRNA; cDNA DKFZp586M0723 (from clone DKFZp586M0723)
AK021798	4.56E-05	0.6377454	FLJ11736 | hypothetical protein FLJ11736
AI268007	4.58E-05	0.7185686	Homo sapiens cDNA FLJ30137 fis, clone BRACE2000078
BC001403	4.70E-05	2.4561451	CPSF5 | cleavage and polyadenylation specific factor 5, 25 kD subunit
AK000081	5.38E-05	2.3154373	CDC2L1 | cell division cycle 2-like 1 (PITSLRE proteins)
BC014270	5.53E-05	2.0457284	PRKCZ | protein kinase C, zeta
AL117478	5.97E-05	1.7598438	AGS3 | likely ortholog of rat activator of G-protein signaling 3
BF116098	7.56E-05	0.4180467	ESTs
BC006499	7.83E-05	1.8287714	HRAS | v-Ha-ras Harvey rat sarcoma viral oncogene homolog
NM_003489	7.94E-05	0.4637752	NRIP1 | nuclear receptor interacting protein 1
AI469557	8.50E-05	1.8599762	EPHB3 | EphB3
AI561249	9.19E-05	0.4329273	KTN1 | kinectin 1 (kinesin receptor)
BC015497	9.45E-05	1.9287915	TEAD4 | TEA domain family member 4

Accession	p value	hazard ratio	Description
AL133661	1.08E-04	0.4897642	DKFZp434C0328 | hypothetical protein DKFZp434C0328
BC015594	1.10E-04	2.0502453	Homo sapiens mRNA for FLJ00083 protein, partial cds
AW135596	1.14E-04	0.6460164	FLJ10058 | hypothetical protein FLJ10058
AI033912	1.18E-04	0.6482864	RLN2 | relaxin 2 (H2)
NM_020978	1.28E-04	0.598655	AMY2B | amylase, alpha 2B; pancreatic
BC006437	1.49E-04	2.0560166	C321D2.4 | hypothetical protein C321D2.4
AW016075	1.51E-04	0.5312489	ESTs, Weakly similar to ALUA_HUMAN !!!! ALU CLASS A WARNING ENTRY !!! [H.sapiens]
NM_001354	1.52E-04	1.4085552	AKR1C2 | aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III)
BC007932	1.56E-04	0.5115812	FLJ11588 | hypothetical protein FLJ11588
AF319520	1.57E-04	1.4189657	ARG99 | ARG99 protein
AA806831	1.62E-04	1.470609	ESTs
AI638324	1.64E-04	0.4669648	Homo sapiens cDNA FLJ30332 fis, clone BRACE2007254
AK025141	1.70E-04	0.6098107	Homo sapiens cDNA: FLJ21488 fis, clone COL05445
AF068918	2.11E-04	1.7571167	BIN1 | bridging integrator 1
AF208111	2.18E-04	0.6637063	IL17BR | interleukin 17B receptor
AK024715	2.34E-04	0.5237823	FLJ21062 | hypothetical protein FLJ21062
BC007836	2.45E-04	1.8806038	MDFI | MyoD family inhibitor
AW192535	2.64E-04	0.46396	ESTs
AA480069	2.68E-04	1.970316	KIAA1925 | KIAA1925 protein
AK025862	2.84E-04	0.4739154	Homo sapiens cDNA: FLJ22209 fis, clone HRC01496
AI800042	2.92E-04	0.4939835	ESTs
AA977269	3.02E-04	1.3578379	FOXD1 | forkhead box D1
BC018644	3.03E-04	1.6098715	NUDT8 | nudix (nucleoside diphosphate linked moiety X)-type motif 8
NM_004419	3.08E-04	0.6155024	DUSP5 | dual specificity phosphatase 5
AW070918	3.10E-04	2.0916912	ESTs, Weakly similar to T2D3_HUMAN TRANSCRIPTION INITIATION FACTOR TFIID 135 KDA SUBUNIT [H.sapiens]
AA040945	3.22E-04	2.2990713	ESTs
AF035282	3.30E-04	0.6524492	C1orf21 | chromosome 1 open reading frame 21
NM_006304	3.34E-04	0.4895086	DSS1 | Deleted in split-hand/split-foot 1 region
R62589	3.47E-04	0.6003814	ESTs
AI400775	3.52E-04	2.2438708	RABL2B | RAB, member of RAS oncogene family-like 2B
AI128331	3.60E-04	0.5099963	ENDOFIN | endosome-associated FYVE-domain protein
AW069725	3.62E-04	0.5812922	CRYZ | crystallin, zeta (quinone reductase)
AK024967	3.82E-04	0.4618762	Homo sapiens cDNA: FLJ21314 fis, clone COL02248
AK022916	3.88E-04	0.5564747	ZNF281 | zinc finger protein 281
BC015484	3.92E-04	1.5502435	CALB2 | calbindin 2, (29kD, calretinin)
AI953054	4.06E-04	1.9805492	TKT | transketolase (Wernicke-Korsakoff syndrome)
BE675157	4.28E-04	0.6073104	ESTs
AF153330	4.33E-04	0.5983906	SLC19A2 | solute carrier family 19 (thiamine transporter),

Accession	p value	hazard ratio	Description
			member 2
AL133427	4.35E-04	0.4914871	Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 261172
BF438928	4.77E-04	0.5752913	ESTs
NM_002428	4.77E-04	1.81811	MMP15 | matrix metalloproteinase 15 (membrane-inserted)
AI264644	4.82E-04	1.8613174	KIAA0775 | KIAA0775 gene product
BE967259	4.88E-04	0.7445998	BCL2 | B-cell CLL/lymphoma 2
AW076080	4.93E-04	0.5435194	Homo sapiens, clone IMAGE:3463399, mRNA, partial cds
T52873	5.05E-04	0.5449457	ESTs, Moderately similar to G02075 transcription repressor zinc finger protein 85 [H.sapiens]
AF085233	5.10E-04	0.635643	SGKL | serum/glucocorticoid regulated kinase-like
BE671445	5.12E-04	0.5796479	ESTs
AI356375	5.23E-04	1.7149531	CDKN2A | cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)
BF589163	5.28E-04	0.5585288	ESTs
AA909006	5.35E-04	1.5526313	LBP-32 | LBP protein 32
BC015792	5.47E-04	1.841097	Homo sapiens, clone MGC:23665 IMAGE:4866941, mRNA, complete cds
BC000692	5.61E-04	2.0170046	HYAL2 | hyaluronoglucosaminidase 2
AL050090	5.73E-04	0.7500215	DKFZP586F1018 | DKFZP586F1018 protein
NM_020235	5.94E-04	0.5893936	BBX | bobby sox homolog (Drosophila)
BF433657	5.99E-04	1.9378811	ESTs
AI692302	6.01E-04	1.899281	ESTs
AK024782	6.05E-04	1.9756718	KIAA1608 | KIAA1608 protein
AF124735	6.12E-04	1.4649329	LHX2 | LIM homeobox protein 2
BC007066	6.12E-04	0.5216856	CDA11 | CDA11 protein
AW135238	6.20E-04	0.4896724	ESTs
AK026747	6.44E-04	0.5015784	LOC54103 | hypothetical protein
AA542898	6.46E-04	0.7842204	P28 | dynein, axonemal, light intermediate polypeptide
BC014913	6.52E-04	0.6913458	Homo sapiens, Similar to synaptotagmin-like 4, clone MGC:17313 IMAGE:3908307, mRNA, complete cds
AI270018	6.72E-04	2.0809844	ECE1 | endothelin converting enzyme 1
L13738	6.90E-04	1.6894154	ACK1 | activated p21cdc42Hs kinase
BC002607	7.01E-04	1.5250234	KIAA1446 | KIAA1446 protein
BI793002	7.18E-04	0.4917655	OSBPL8 | oxysterol binding protein-like 8
BC007092	7.20E-04	1.2827239	HOXB13 | homeo box B13
BC009874	7.40E-04	1.730815	JUN | v-jun sarcoma virus 17 oncogene homolog (avian)
AF321193	7.41E-04	1.5356899	DSCR8 | Down syndrome critical region gene 8
AK000397	7.70E-04	1.5631718	FLJ10351 | likely ortholog of mouse piwi like homolog 1 (Drosophila)-like
AF052110	7.76E-04	1.6400255	DAF | decay accelerating factor for complement (CD55, Cromer blood group system)
AA648536	8.03E-04	1.6290887	MYO1E | myosin IE
BF436400	8.31E-04	0.7911405	EST

Accession	p value	hazard ratio	Description
AL050179	8.59E-04	0.5180149	TPM1 | tropomyosin 1 (alpha)
AI700363	8.60E-04	1.3675668	ESTs
NM_004038	8.72E-04	0.6247207	AMY1A | amylase, alpha 1A; salivary
AF060555	8.75E-04	1.5560891	ESR2 | estrogen receptor 2 (ER beta)
AK026756	8.85E-04	0.6360787	KIAA1603 | KIAA1603 protein
AI686003	8.97E-04	0.6087104	ESTs
NM_019120	9.14E-04	1.4302118	PCDHB8 | protocadherin beta 8
NM_020957	9.50E-04	1.4881037	PCDHB16 | protocadherin beta 16
AI921700	9.73E-04	0.522736	ITGAV | integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51)
X62534	9.87E-04	0.5796731	HMG2 | high-mobility group (nonhistone chromosomal) protein 2
BC002738	9.90E-04	1.8608522	CRIP1 | cysteine-rich protein 1 (intestinal)

Between the two approaches, 114 genes were in common. At the significance level of 0.001, about 6 genes are expected by chance if there are no real differences between the patient groups, indicating that the 149 genes identified by either method are highly statistically significant.

Example 2

Kaplan-Meier survival curves of patients stratified by cross-validation

Kaplan-Meier analysis was performed to assess the differential survival of patients stratified by the gene expression signature. Leave-one-out-cross-validation was performed. Briefly, one of the 62 patients was left out as a test sample, and the other 61 samples were used in Cox regression to both select significant genes ($p < 0.001$) and obtain gene-specific weights (Cox regression coefficients β). A linear sum of the gene-specific weights (β) times expression levels (x) across all selected genes was calculated as the overall risk score for each patient: $S = \text{sum}(\beta_i x_i)$ for all selected genes. The mid-point $\boldsymbol{m}$ between the median scores for the two patient groups (recurrence/non-recurrence) in the training set was calculated: $\boldsymbol{m}$ = (median score of recurrence group + median score of non-recurrence group)/ 2, and the score for the test sample $\boldsymbol{S}$ was compared with $\boldsymbol{m}$ to classify the test sample to either the recurrence ($\boldsymbol{S} > \boldsymbol{m}$, TAM signature-) or non-recurrence group ($\boldsymbol{S} <= \boldsymbol{m}$, TAM signature+). This entire procedure was repeated 62 times to

generate a classification for each patient. Disease-free survival curves of the two groups as assigned by the cross-validation procedure are compared. The results are shown in Figure 1.

Example 3

Identification of biomarker predictors of TAM treatment outcome

Samples from 60 patients with ER+ primary breast cancer, and treated with adjuvant TAM, were selected tamoxifen based on treatment outcome. 28 had developed tumor recurrence with a median time of 4 years, and 32 remained disease-free with a median follow-up of 10 years (Table 3). Patients who remained disease-free during the entire follow up period were likely to represent responders to TAM, although a small subset of them might have been cured by surgery alone. Those patients who developed tumor recurrence despite TAM therapy either did not respond or developed resistance to TAM and are hereafter referred to as non-responders for brevity. To control for known prognostic factors, tumors between these two groups were matched by tumor size, lymph node status and tumor grade.

Table 3 Patients and tumor characteristics

Sample ID	Tumor type	Size	Grade	Nodes	ER	PR	Age	DFS	Status
1389	D	1.7	2	0/1	Pos	Pos	80	94	0
648	D	1.1	2	0/15	Pos	ND	62	160	0
289	D	3	2	0/15	Pos	ND	75	63	1
749	D	1.8	2	2/9	Pos	Pos	61	137	0
420	D/L	2	3	ND	Pos	Pos	72	58	1
633	D	2.7	3	0/11	Pos	ND	61	20	1
662	D	1	3	6/11	Pos	Pos	79	27	1
849	D	2	1	0/26	Pos	Neg	75	23	1
356	D	1	2	2/20	Pos	ND	58	24	1
1304	D	2	3	0/14	Pos	Pos	57	20	1
1419	D	2.5	2	1/8	Pos	Pos	59	86.04	0
1093	D	1	3	1/14	Pos	Pos	66	84.96	0
1047	D/L	2.6	2	0/18	Pos	Neg	70	127.92	0
1037	D/L	1.5	2	0/4	Pos	Pos	85	83.04	0
319	D	4	2	1/13	Pos	ND	67	44	1
25	D	3.5	2	0/9	Neg	Pos	62	75	1
180	D	1.6	2	2/19	Pos	Pos	69	168.96	0
687	D	3.5	3	3/16	Pos	ND	73	141.96	0

856	D	1.6	2	0/16	Pos	Pos	73	87.96	0
1045	D	2.5	3	1/12	Pos	Neg	73	120.96	0
1205	D	2.7	2	1/19	Pos	Pos	71	87.96	0
1437	D	1.7	2	2/22	Pos	Pos	67	89.04	0
1507	D	3.7	3	0/40	Pos	Pos	70	69.96	0
469	D	1	1	0/19	Pos	ND	66	161.04	0
829	D	1.2	2	0/9	Pos	ND	69	135.96	0
868	D	3	3	0/13	Pos	Pos	65	129.96	0
1206	D	4.1	3	0/15	Pos	Neg	84	56	1
843	D	3.4	2	11/20	Pos	Neg	76	122	1
342	D	3	2	9/21	Pos	ND	62	102	1
1218	D	4.5	1	3/16	Pos	Pos	62	10	1
547	D/L	1.5	2	ND	Pos	ND	74	129	1
1125	D	2.6	2	0/18	Pos	Pos	54	123	0
1368	D	2.6	2	ND	Pos	Pos	82	63	0
605	D	2.2	2	6/18	Pos	ND	70	110.04	0
59	L	3	2	33/38	Pos	ND	70	21	1
68	D	3	2	0/17	Pos	ND	53	38	1
317	D	1.2	3	1/10	Pos	Pos	71	5	1
374	D	1	3	0/15	Pos	Neg	57	47	1
823	D	2	2	0/6	Pos	Pos	51	69	1
280	D	2.2	3	0/12	Pos	ND	66	44	1
651	D	4.7	3	10/13	Pos	ND	48	137	1
763	D	1.8	2	0/14	Pos	Pos	63	117.96	0
1085	D	4.7	2	0/8	Pos	Pos	48	101	1
1363	D	2.1	2	0/15	Pos	Pos	56	114	0
295	D	3.5	2	3/21	Pos	Pos	52	118	1
871	D	4	3	0/16	Pos	Neg	61	6	1
1343	D	2.5	3	ND	Pos	Pos	79	21	1
140	L	>2.0	2	18/28	Pos	ND	63	43	1
260	D/L	0.9	2	1/13	Pos	ND	73	42	1
297	D	0.8	2	1/16	Pos	Pos	66	169	0
1260	D	3.5	2	0/14	Pos	Pos	58	79	0
1405	D	1	3	ND	Pos	Pos	81	95.04	0
518	L	5.5	2	3/20	Pos	ND	68	156	0
607	D	1.2	2	5/14	Pos	Pos	76	114	0
638	D	2	2	1/24	Pos	Pos	67	147.96	0
655	D	2	3	ND	Pos	Pos	73	143.04	0
772	D	2.5	2	0/18	Pos	Pos	68	69	1
878	D/L	1.6	2	0/9	Pos	Neg	76	138	0
1279	D	2	2	0/12	Pos	Pos	68	102	0
1370	D	2	2	ND	Pos	Pos	73	60.96	0

Abbreviations: D, ductal; L, lobular; pos, positive; neg, negative; ND, not determined; ER, estrogen receptor; PR, progesterone receptor; DFS, disease-free survival; status=1, recurred; status=0, disease-free.

The samples were used to identify gene expression signatures correlated with outcome of TAM treatment. Each breast cancer biopsy contains a mixture of cell types including epithelial breast cancer cells, infiltrating lymphocytes, endothelial cells and stromal fibroblasts. It has been suggested that complex interactions among these cell types in the tumor microenvironment determine the biological behavior of the tumor. Therefore, to identify gene expression differences in primary tumors between TAM responders and non-responders, expression profiling of both whole tissue sections, which represent this microenvironment, and microdissected, largely pure populations of epithelial cancer cells from each tumor biopsy were conducted on a custom 22k oligonucleotide microarray.

This generated two parallel datasets corresponding to each patient: one set from whole tissue sections ("sections dataset") and another from laser capture microdissected cancer cells ("LCM dataset"). Each expression dataset was first filtered based on overall variance of each gene and the top 5475 high-variance genes (75th percentile) were selected. Using the reduced datasets, t-test on each gene between the TAM responders and non-responders were carried out. From the sections dataset, 19 genes were identified at the p value cutoff of 0.001 (Table 4). The probability of selecting this many or more differentially expressed genes by chance was 0.035 as estimated by randomly permuting the patient class with respect to treatment outcome and repeating the t-test procedure 1000 times. Among the 19 genes identified in the sections dataset, genes involved in immune response are particularly prominent.

Table 4. 19-gene signature identified by t-test in the Sections dataset

	Parametric p-value	Mean in responders	Mean in non-responders	Fold difference of means	GB acc	Description
1	1.96E-05	0.759	1.317	0.576	AW006861	SCYA4 | small inducible cytokine A4
2	2.43E-05	1.31	0.704	1.861	AI240933	ESTs
3	8.08E-05	0.768	1.424	0.539	X59770	IL1R2 | interleukin 1 receptor, type II
4	9.57E-05	0.883	1.425	0.62	AB000520	APS | adaptor protein with pleckstrin homology and src homology 2 domains
5	9.91E-05	1.704	0.659	2.586	AF208111	IL17BR | interleukin 17B receptor
6	0.0001833	0.831	1.33	0.625	AI820604	ESTs
7	0.0001935	0.853	1.459	0.585	AI087057	DOK2 | docking protein 2, 56kD

8	0.0001959	1.29	0.641	2.012	AJ272267	CHDH | choline dehydrogenase
9	0.0002218	1.801	0.943	1.91	N30081	ESTs, Weakly similar to I38022 hypothetical protein [H.sapiens]
10	0.0004234	1.055	2.443	0.432	AI700363	ESTs
11	0.0004357	0.451	1.57	0.287	AL117406	ABCC11 | ATP-binding cassette, sub-family C (CFTR/MRP), member 11
12	0.0004372	1.12	3.702	0.303	BC007092	HOXB13 | homeo box B13
13	0.0005436	0.754	1.613	0.467	M92432	GUCY2D | guanylate cyclase 2D, membrane (retina-specific)
14	0.0005859	1.315	0.578	2.275	AL050227	Homo sapiens mRNA; cDNA DKFZp586M0723 (from clone DKFZp586M0723)
15	0.000635	1.382	0.576	2.399	AW613732	Homo sapiens cDNA FLJ31137 fis, clone IMR322001049
16	0.0008714	0.794	1.252	0.634	BC007783	SCYA3 | small inducible cytokine A3
17	0.0008912	2.572	1.033	2.49	X81896	C11orf25 | chromosome 11 open reading frame 25
18	0.0009108	0.939	1.913	0.491	BC004960	MGC10955 | hypothetical protein MGC10955
19	0.0009924	1.145	0.719	1.592	AK027250	Homo sapiens cDNA: FLJ23597 fis, clone LNG15281

Repeating the same analysis on the LCM dataset yielded 9 significant genes at the cutoff of $p < 0.001$ (Table 5); however, the probability of finding 9 or more genes by chance is 0.154 in permutation analysis (n=1000). These results established that significant differences in gene expression between the two patient groups exist, but differences were subtle.

Table 5. 9-gene signature identified by t-test in the LCM dataset

	Parametric p-value	Mean in responders	Mean in non-responders	Fold difference of means	GB acc	Description
1	2.67E-05	1.101	4.891	0.225	BC007092	HOXB13 | homeo box B13
2	0.0003393	1.045	2.607	0.401	AI700363	ESTs
3	0.0003736	0.64	1.414	0.453	NM_014298	QPRT | quinolinate phosphoribosyltransferase (nicotinate-nucleotide pyrophosphorylase (carboxylating))
4	0.0003777	1.642	0.694	2.366	AF208111	IL17BR | interleukin 17B receptor
5	0.0003895	0.631	1.651	0.382	AF033199	ZNF204 | zinc finger protein 204
6	0.0004524	1.97	0.576	3.42	AI688494	FLJ13189 | hypothetical protein FLJ13189
7	0.0005329	1.178	0.694	1.697	AI240933	ESTs
8	0.0007403	0.99	1.671	0.592	AL157459	Homo sapiens mRNA; cDNA DKFZp434B0425 (from clone DKFZp434B0425)

9	0.0007739	0.723	1.228	0.589	BC002480	FLJ13352 | hypothetical protein FLJ13352

The sequence of each GenBank accession number in Tables 4 and 5 is presented in the attached Appendix.

Due to the limited sample size (n=60), leave-one-out cross validation was used to assess the predictive significance of the gene expression signature. In each round of cross validation, significant genes were identified using the training set by t-test at $p < 0.001$, and a compound covariate predictor was built as the linear combination oft he gene expression values over all significant genes weighted by their t-statistics. The predictor was then used to predict the left-out sample. Repeating this procedure 60 times generated an "honest" prediction on each sample.

Using the sections dataset, the overall accuracy of cross validation results are 70%, and the sensitivity, specificity, positive and negative predictive values are 60%, 78%, 71%, and 69%, respectively. The results of analyzing the LCM dataset were slightly lower, with an overall accuracy of 67%, and sensitivity, specificity, positive and negative predictive values of 57%, 75%, 67%, and 67%, respectively. Patients having the "responder signature" and those having the "non-responder signature" as predicted from cross validation demonstrate significantly different disease-free survival curves (Fig. 2).

Previously a 70-gene prognostic classifier was derived from correlating gene expression profiles with distant metastasis from node-negative breast cancer patients, most of which received no adjuvant chemotherapy or endocrine therapy. 61 of the 70 genes from the study were on the microarrays used in this example. Expression data corresponding to these 61 genes were extracted from the sections dataset because the 70-gene signature study used whole tissue sections. None of these 61 genes were significantly differentially expressed between TAM responders and non-responders at the significance level of 0.001, and only 3 genes were significant at $p < 0.05$. Leave-one-out cross-validation analysis using either all 61 genes or only genes with $p < 0.05$ gave overall accuracies of 52% and 53% respectively. Thus the 70-gene classifier derived from mostly untreated patients cannot predict tumor recurrence after adjuvant TAM treatment. Without being bound by theory, and offered to improve the understanding of the invention, this suggests that the treatment

outcome by TAM is not simply a reflection of the aggressiveness of the primary tumor, but may directly reflect the responsiveness to TAM.

Example 4

Identification of 3 biomarker predictors of TAM treatment outcome

Between the two sets of significant genes identified with the sections and LCM datasets of Example 3, 4 genes (AI700363, EST; BC007092, HOXB13; AF208111, IL17BR; AI240933, EST) were in common. Further sequence analysis indicated that the EST sequence AI700363 represents a splicing variant of HOXB13 and the other EST (AI240933) represents the 3' end of the putative calcium channel gene CACNA1D. Therefore, these analyses identified three distinct genes having statistically significant differential expression between responders and non-responders (Fig. 3). It is noteworthy that HOXB13 had a more significant difference between responders and non-responders in the LCM dataset. The fact that these three genes were identified both in the sections and LCM datasets serves to validate the microarray measurements, and also suggest that they are likely to be differentially expressed by the tumor cells themselves.

The significant correlations of CACNA1D, HOXB13 and IL17BR with TAM treatment outcome suggest that these three genes may be novel predictors of TAM response. Estrogen receptor status is a powerful predictor of response to tamoxifen, as 60% ER+ vs. < 10% ER- tumors respond to TAM. However, among ER+ tumors, no established predictors exist to identify the 40% non-responders. Therefore, the predictive usefulness of CACNA1D, HOXB13 and IL17BR as potential biomarkers to identify the ER+, TAM responders and non-responders was tested.

Receiver operating characteristic (ROC) analysis evaluates the sensitivity and specificity of a clinical test. The area under the curve (AUC) of plotting the false positive rate against the true positive rate measures the overall accuracy. In both the sections and LCM datasets, all three genes demonstrated consistent predictive ROC curves (Fig. 3). The AUC values (Table 4) for IL17BR and CACNA1D ranges from 0.76 to 0.81 with higher values in the sections data; HOXB13 has considerably higher AUC in the LCM dataset than in the sections dataset (0.79 vs. 0.69), consistent with the t-test results (Fig. 4). Statistical test for AUC > 0.5 indicates that all AUC values are significant (Table 6).

Table 6. ROC analysis summary

	Sections		LCM		FFPE	
	AUC	*P*	AUC	*P*	AUC	*P*
IL17BR	0.79	1.58E-06	0.76	2.73E-05	0.83	4.94E-06
CACNA1D	0.81	3.02E-08	0.76	1.59E-05	0.79	1.54E-04
HOXB13	0.67	0.012	0.79	9.94E-07	0.58	0.216
ESR1	0.55	0.277	0.63	0.038	0.58	0.218
PGR	0.65	0.020	0.63	0.039	0.58	0.247
ERBB2	0.69	0.004	0.64	0.027	0.59	0.226
EGFR	0.56	0.200	0.61	0.068	0.62	0.133

AUC, area under the curve; *P* values compare AUC > 0.5.

As a further demonstration for the predictive utility of CACNA1D, HOXB13 and IL17BR, Kaplan-Meier analysis was performed to assess the correlation of the expression levels with disease-free survival. For each gene, patients were stratified into two groups using the median as cut point: low (<= median) and high (> median), and the Kaplan-Meier curves were compared in log-rank test (Fig. 5). Stratification by each of these three genes results in two groups with highly significant different disease-free survival times.

Considerable evidence suggests that the activity of growth factor signaling pathways may negatively regulate estrogen signaling, which may contribute to loss of responsiveness or developing resistance to TAM. Therefore, we evaluated the predictive utility of ESR1, PGR (positive predictors), ERBB2 and EGFR (negative predictors) in our datasets by ROC analysis. The AUCs ranged from 0.55 to 0.69 for these genes, but the values of PGR and ERBB2 were significantly higher than 0.5 in both sections and LCM datasets (Table 6), which is consistent with prior studies. Taken together, these results demonstrate that the three genes identified in this study are significantly stronger than estrogen and progesterone receptors as positive predictors and ERBB2 and EGFR as negative predictors.

We next validated these results using an independent cohort of 31 patients uniformly treated with TAM. Primary breast cancer biopsies in the form of formalin-fixed paraffin-embedded (FFPE) blocks were used for microarray analysis; macro-dissection was performed to enrich for tumor content. The expression levels of CACNA1D, HOXB13, and IL17BR were compared between

responders (n=9) and non-responders (n=22) (Fig. 6) and ROC analysis performed as before (Fig. 6; Table 6). The three genes showed statistically significant differences in gene expression between TAM responders and non-responders similar to those seen in the sections and LCM datasets (Fig. 6, cf. Figs. 3-4). The AUC values for IL17BR and CACNA1D are 0.83 and 0.79, respectively; AUC for HOXB13 was insignificant but with a consistent trend in the earlier portions of the ROC curve. Compared to the known genes (ESR1, PGR, ERBB2 and EGFR), IL17BR and CACNA1D were significantly stronger predictors of TAM response (Table 6).

Because HOXB13 and IL17BR display opposing patterns of expression, the idea of using the ratio of HOXB13 over IL17BR as a composite predictor was tested (Fig. 7). Two sample t-tests indicated that the two-gene ratio had a stronger correlation with treatment outcome than either gene alone in both the sections and FFPE datasets (Fig. 7; cf. Fig, 3). ROC curves have AUCs of 0.8 and 0.83 for the sections and FFPE data, respectively. From the ROC curve for the sections data, minimizing the absolute difference between sensitivity and specificity yielded an optimal cut point of –0.22 (log2 scale) (horizontal line in Fig. 7). Classifying the patients in the sections data into responders (log ratio <= -0.22) and non-responders (log ratio > -0.22) resulted in correct classification of 46 of the 60 patients (77%, p=4.224e-05, 95% CI 64%-87%). Applying the same classification rule to the FFPE dataset, 8 of the 9 responders and 16 of the 22 non-responders were correctly classified (overall accuracy = 77%, p-value = 0.003327, 95% CI 59%-90%).

Example 5

Multivariate analysis

Expression data from the three genes were used in logistic regression models by calculating cross-validated compound covariate scores as linear combinations of the expression values of the three genes weighted by their t-test statistics in each round of leave-one-out cross validation. The compound covariate score has a univariate p value of 0.0003 with both sections and LCM datasets, and the model had a bootstrap-adjusted accuracy of 81% (Table 7). Next, multivariate logistic regression analysis was performed using clinicopathological factors plus the compound covariate score. Because only two samples were grade 1, grades 1 and 2 were combined into one level (low-grade) and compared to grade 3 (high-grade). Due to missing values in clinical parameters, 53

cases were used for modeling. The multivariate model shows that the compound covariate score was the only independent significant predictor (Table 7). Clinical factors (such as tumor size, grade and nodal status) were not significantly associated with TAM treatment outcome.

Table 7. Multivariate analysis

PREDICTIVE POWER OF BREAST CANCER RECURRENCE OF EACH INDIVIDUAL PREDICTOR [1]

Model 1:

	LCM DATA				SECTION DATA			
Accuracy [2]	0.807				0.817			
Predictors	Odds Ratio	Lower 95% CI of Odds Ratio	Upper 95% CI of Odds Ratio	P Value	Odds Ratio	Lower 95% CI of Odds Ratio	Upper 95% CI of Odds Ratio	P Value
Score of Genes [3]	7.4	2.5	21.8	0.0003	8.7	2.7	28.2	0.0003

Model 2:

	LCM DATA				SECTION DATA			
Accuracy [2]	0.796				0.798			
Predictor	Odds Ratio	Lower 95% CI of Odds Ratio	Upper 95% CI of Odds Ratio	P Value	Odds Ratio	Lower 95% CI of Odds Ratio	Upper 95% CI of Odds Ratio	P Value
Tumor size	1.2	0.5	3.0	0.662	1.3	0.6	3.1	0.544
Nodal status (pos:neg)	0.8	0.2	3.2	0.777	0.9	0.2	3.4	0.840
Tumor grade (high:low)	1.5	0.3	6.5	0.619	1.2	0.3	5.9	0.793
Score of Genes [3]	8.5	2.2	33.3	0.0021	10.8	2.4	48.0	0.0018

[1] Model P value is estimated based upon a multivariate logistic regression model against tumor recurrence status.
[2] Model predictive accuracy is estimated based on bias-adjusted AUC index by 200 bootstraps.
[3] Score of genes is a pre-validated compound covariance score based on gene expressions levels and the regression coefficient for each predictor based on univariate logistic regression model.

The results reflected in Table 7 are expected because the responder and non-responder groups were matched by these parameters in patient selection. Bootstrap validation analysis indicated that the full model has a concordance index of 80%. Taken together, these results

demonstrate that the three genes identified in this study were strong independent predictors of treatment outcome by adjuvant therapy independent of known clinicopathological parameters.

All references cited herein, including patents, patent applications, and publications, are hereby incorporated by reference in their entireties, whether previously specifically incorporated or not.

Having now fully described this invention, it will be appreciated by those skilled in the art that the same can be performed within a wide range of equivalent parameters, concentrations, and conditions without departing from the spirit and scope of the invention and without undue experimentation.

While this invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications. This application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth.

Appendix

Sequences identified as those of IL17RB cluster

AW675096

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGA
GAGCCGACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACAT
GATCTAATCCCGGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACTAGTGTTGCA
ACAGGGGACTATTCAATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATC
CGCTTGTTGAAGGCCACCAAGATTTGTGTGACGGGCAAAAGCAACTTCCAGTCCTACAGC
TGTGTGAGGTGCAATTACACAGAGGCCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAA
TGGACATTTTCCTACATCGGCTTCCCTGTAGAGCTGAACACAGTCTATTTCATTGGGGCC
CATAATATTCCTAATGCAAATATGAATGAAGATGGCCCTTCCATGTCTGTGAATNTCACC
TCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAAAAGTGTGTCAAGGCCGGAAGC
CTGTGGGATCCGAACATCACT

AW673932

TTTTTTTTTTTTTTTTTTTAAAAGTGGGTTCAGCTTGTTTATTCCCTACTTTTGTTATCT
TAAAAACAATGATTTTTTGCATGTAATAGAAGGTTTTTCACTTAAGATGCTATTGAGTGA
ATCAGTGAGGGGTTCTTAGAGTTAGTATTCATTAATTAAACATAGAATATTAGCTAAACA
GTTCTGGGTACACTGCAATGCATGGTCTATGGAAGACTAGATGTTTGGCTGAAGATGCTT
TATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAACTGTAATTGATTTCTATGT
ATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTAGTTAATGCTACATTAGTT
AGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAAGGCTGTTTGTAGGC
TGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATTGG

BC000980

ggcccggcga tgtcgctcgt gctgctaagc ctggccgcgc tgtgcaggag cgccgtaccc
cgagagccga ccgttcaatg tggctctgaa actgggccat ctccagagtg gatgctacaa
catgatctaa tcccgggaga cttgagggac ctccgagtag aacctgttac aactagtgtt
gcaacagggg actattcaat tttgatgaat gtaagctggg tactccgggc agatgccagc
atccgcttgt tgaaggccac caagatttgt gtgacgggca aaagcaactt ccagtcctac
agctgtgtga ggtgcaatta cacagaggcc ttccagactc agaccagacc ctctggtggt
aaatggacat tttcctacat cggcttccct gtagagctga acacagtcta tttcattggg
gcccataata ttcctaatgc aaatatgaat gaagatggcc cttccatgtc tgtgaatttc
acctcaccag gctgcctaga ccacataatg aaatataaaa aaaagtgtgt caaggccgga
agcctgtggg atccgaacat cactgcttgt aagaagaatg aggagacagt agaagtgaac
ttcacaacca ctcccctggg aaacagatac atggctctta tccaacacag cactatcatc
gggttttctc aggtgtttga gccacaccag aagaaacaaa cgcgagcttc agtggtgatt
ccagtgactg gggatagtga aggtgctacg gtgcagctga ctccatattt tcctacttgt

ggcagcgact gcatccgaca taaaggaaca gttgtgctct gcccacaaac aggcgtccct
ttccctctgg ataacaacaa aagcaagccg ggaggctggc tgcctctcct cctgctgtct
ctgctggtgg ccacatgggt gctggtggca gggatctatc taatgtggag gcacgaaagg
atcaagaaga cttccttttc taccaccaca ctactgcccc ccattaaggt tcttgtggtt
tacccatctg aaatatgttt ccatcacaca atttgttact tcactgaatt tcttcaaaac
cattgcagaa gtgaggtcat ccttgaaaag tggcagaaaa agaaaatagc agagatgggt
ccagtgcagt ggcttgccac tcaaaagaag gcagcagaca aagtcgtctt ccttctttcc
aatgacgtca acagtgtgtg cgatggtacc tgtggcaaga gcgagggcag tcccagtgag
aactctcaag acctcttccc ccttgccttt aacctttct gcagtgatct aagaagccag
attcatctgc acaaatacgt ggtggtctac tttagagaga ttgatacaaa agacgattac
aatgctctca gtgtctgccc caagtaccac ctcatgaagg atgccactgc tttctgtgca
gaacttctcc atgtcaagca gcaggtgtca gcaggaaaaa gatcacaagc ctgccacgat
ggctgctgct ccttgtagcc cacccatgag aagcaagaga ccttaaaggc ttcctatccc
accaattaca gggaaaaaac gtgtgatgat cctgaagctt actatgcagc ctacaaacag
ccttagtaat taaaacattt tataccaata aaattttcaa atattgctaa ctaatgtagc
attaactaac gattggaaac tacatttaca acttcaaagc tgttttatac atagaaatca
attacagttt taattgaaaa ctataaccat tttgataatg caacaataaa gcatcttcag
ccaaacatct agtcttccat agaccatgca ttgcagtgta cccagaactg tttagctaat
attctatgtt taattaatga atactaactc taagaacccc tcactgattc actcaatagc
atcttaagtg aaaaaccttc tattacatgc aaaaaatcat tgtttttaag ataacaaaag
tagggaataa acaagctgaa cccactttta aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
aa

BI602183

AGCGGAGCTGCGGGTGGCCTGGATCCCGCGCAGTGGCCCGGCGATGTCGCTCGTGCTGCT
AAGCCTGGCCACGCTGTGCAGGAGCGCCGTACCCCGAGAGCCGACCGTTCAATGTGGCTC
TGAAACTGTGGACATTTTCCTATATCGGCTTCCCTGTAGAGCTGAAAACAGTCTATTTCA
TTGGGGCCCATAATATTCCTAATGCAAATATGAATGAAGATGGCCCTTCCATGTCTGTGA
ATTTCACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAAAGTGTGTCAAGGC
CGGAAGCCTGTGGGATCCGAACATCACTGCTTGTAAGAAGAATGAGGAGACAGTAGAAGT
GAACTTCACAACCACTCCCCTGGGAAACAGATACATGGCTCATCCAACACAGCACTATCA
TCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCCTACTT
GTGGCAGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCC
CTTTCCCCTCTGGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTG
TCTCTGCTGGTTGGCCACATTGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACG
AAAGGATCCAGAAGACTTCCTTTTCTACCACAAACTACTGCCCCCATTAAGGTCCTGTGG
TTACCCATCTTGAAATATGTTCCTCACACAATTTGTTACTTCACTGAATTCTTCAAAACC
TG

BI458542

AGCGGAGCGTGCGGGTGGCCTGGATCCCGCGCAGTGGCCCGGCGATGTCGCTCGTGCTGC

TAAGCCTGGCCACGCTGTGCAGGAGCGCCGTACCCCGAGAGCCGACCGTTCAATGTGGCT
CTGAAACTGTGGACATTTTCCTATATCGGCTTCCCTGTAGAGCTGAAAACAGTCTATTTC
ATTGGGGCCCATAATATTCCTAATGCAAATATGAATGAAGATGGCCCTTCCATGTCTGTG
AATTTCACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAAAAGTGTGTCAAG
GCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTAAGAAGAATGAGGAGACAGTAGAA
GTGAACTTCACAACCACTCCCCTGGGAAACAGATACATGGCTCATCCAACACAGCACTAT
CATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGT
GATTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCCTAC
TTGTGGCAGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGT
CCCTTTCCCTCTGGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCT
GTCTCTGCTGGTGGNCACATTGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACG
AAAGGATCAGAAGACTTCCTTTTCTACCACCACATACTGCCCCCCATTAAGGTTCTTGTG
GTTTACCC

BI823321

GGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGA
GCCGACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGA
TCTAATCCCGGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACTAGTGTTGCAAC
AGGGGACTATTCAATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCG
CTTGTTGAAGGCCACCAAGATTTGTGTGACGGGCAAAAGCAACTTCCAGTCCTACAGCTG
TGTGAGGTGCAATTACACAGAGGCCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATG
GACATTTTCCTATATCGGCTTCCCTGTAGAGCTGAACACAGTCTATTTCATTGGGGCCCA
TAATATTCCTAATGCAAATATGAATGAAGATGGCCCTTCCATGTCTGTGAATTTCACCTC
ACCAGGAAGCCTGTGGGATCCGAACATCACTGCTTGTAAGAAAGAATGAGGAGACAGTAG
AAGTGAACTTCACAACCACTCCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCA
CTATCATCGGGTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGT
GGTGATTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCC
TACTTGTGGCAGCGACTGCAATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAG
GCGTCCCTTTCCCTCTTGGATAGCAACAGAAGCAAGCCGGGAGGCTGGTGCCTCTTCTTC
TGGTGTCTCTGCTGGTGGCACATTGAGTGCTGGTGGCAGGATCCATCTAATGTGGAGGCC
CCAAAGGACCAGGAAAGACTTCCTTTATTAGCACCAAGTATTGCCC

AA514396

TGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAACTGT
AATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTAGTT
AATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTA
AGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATT
GGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCA
GCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTACTTGACATGGAGAAG
TTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGC
ATTGTAATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCACCACGTATTTGTGCAGATG
AATCTGGC

BF110326

TTTGTTTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAA
AACTGTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCG
TTAGTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAA
TTACTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCT
GTAATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTTTCATGGGTGGGCTACAAGGA
GCAGCAGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATG
GAGAAGTTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACT
GAGAGCATTGTAATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCACCACGTATTTGTG
CAGATGAATCTGGCTTCTTAGATCACTGC

BE466508

TGGCATGAGATGCTATATTGTTGCATTATCAAAATGGGTTTAGTCTTCAATTAACACTGT
AATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTGGTTTCCAATCGTCAGTT
AATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTA
AGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATT
GGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCA
GCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAAG
TTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGC
ATTGTAATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCACCACGTATTTGTGCAGATG
AATCTGGCTTCTTAGATCACTG

BF740045

GTTTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAAC
TGTAATTGATTTCTATGTATAAAACACGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTT
AGTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATT
ACTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGT
AATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGC
AGCAGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGA
GAAGTTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGA
GAGCATTGTAATCGTCTTTTGTATCAATCTCTCTAAAGTA

AW299271

TGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAACTGT
AATTGATTTCTATGTATAAAACAGCGTTGAAGTTGTAAATGTAGTTTCCAATCGTTAGTT
AATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTA
AGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATT
GGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCA
GCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAAG

TTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGC
ATTGTAATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCACCACGTATTTGTGCAGATG
AATCTGGCTTCTTAGATCACTGCAGAAAAG

AA836217

TTTTTTTTTTACAACTTCAAAGCTGTTTTATACATAGAAATCAATTACAGTTTTAATTGA
AAACTATAACCATTTTGATAATGCAACAATAAAGCATCTTCAGCCAAACATCTAGTCTTC
CATAGACCATGCATTGCAGTGTACCCAGAACTGTTTAGCTAATATTCTATGTTTAATTAA
TGAATACTAACTCTAAGAACCCCTCACTGATTCACTCAATAGCATCTTAAGTGAAAAACC
TTCTATTACATGCAAAAAATCATTGTTTTTAAGATAACAAAAGTAGGGAATAAACAAGCT
GAACCCACTTTTACTGGACCAAATGATCTATTATATGTGTACCACTTGTATGATTTGGTA
TTTGCATAAGACCTTCCCTCTACAAACTAGATTCATATCTTGATTCTTGTACAGGTGCCT
TTTAACATGAACAACAAAATACCCACAAACTTGTCTACTTTTGCC

AI203628

TAGTAATTAAAACATTTTATACCAATAAAATTTTCAAATATTGCTAACTAATGTAGCATT
AACTAACGATTGGAAACTACATTTACAACTTCAAAGCTGTTTTATACATAGAAATCAATT
ACAGTTTTAATTGAAAACTATAACCATTTTGATAATGCAACAATAAAGCATCTTCAGCCA
AACATCTAGTCTTCCATAGACCATGCATTGCAGTGTACCCAGAACTGTTTAGCTAATATT
CTATGTTTAATTAATGAATACTAACTCTAAGAACCCCTCACTGATTCACTCAATAGCATC
TTAAGTGAAAAACCTTCTATTACATGCAAAAAATCATTGTTTTTAAGATAACAAAAGTAG
GGAATAAACAAGCTGAACCCACTTTTACTGGACCAAATGATCTATTATATGTGTAACCAC
TTGTATGATTTGGTATTTGCATAAGACCTTCCCTCTACAAACTAGATTCATATCTTGATT
CTTGTACAGGTGCCTTTTAACATGAA

AI627783

TTTTTTTTTTTTTTTTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACT
AAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAAT
TGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGC
AGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTACTTGACATGGAGAA
GTTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAG
CATTGTAATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCACCACGTATTTGTGCAGAT
GAATCTGGCTTCTTAGATCACTGCAGAAAAGGTTAAAGGCAAGGGGGAAGAGGTCTTGAG
AGTTCTC

AI744263

TTAAAGTGGGTTCAGCTTGTTTATTCCCTACTTTTGTTATCTTAAAAACAATGATTTTTT
GCATGTAATAGAAGGTTTTTCACTTAAGATGCTATTGAGTGAATCAGTGAGGGGTTCTTA
GAGTTAGTATTCATTAATTAAACATAGAATATTAGCTAAACAGTTCTGGGTACACTGCAA
TGCATGGTCTATGGAAGACTAGATGTTTGGCTGAAGATGCTTTATTGTTGCATTATCAAA

ATGGTTACAGTTTTCAATTAAAGCTGTAATTGATTTCTATGTATAAAACAGCTTTGAAGT
TGTAAATGTAGTTTCCAATCGTTAGTTAATGCTACATTAGTTAGCAATATTTGAAAATTT
TATTGGTATAAAATGTTTTAATTACTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGG
ATCATCACACGTTNTTTCCCTGTAATTGGTGGGATAGGAAGCCTTTA

AI401622

AGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAAGGCTGTTTGT
AGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATTGGTGGGATAGG
AAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCAGCCATCGTGGC
AGGCTTGTGATCTTTTTCCTGCTGACACCTGCTACTTGACATGGAGAAGTTCTGCACAGA
AAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGCATTGTAATCGT
CTTTTGTATCAATCTCTCTAAAGTAGACCACCACGTATTTGTGCAGATGAATCTGGCTTC
TTAGATCACTGCAGAAAAGGTTAAAGGCAAGGGGGAAGAGGTCTTGAGAGTTCTCACTGG

AI826949

TTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAACTG
TAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTAGT
TAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACT
AAGGCTGTTTGTAGGCTTGCATAGAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAAT
TGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGC
AGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAA
GTTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAG
CATTGTAATCGTCT

BE047352

TTTTTTTTTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAAGGCT
GTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATTGGTGG
GATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCAGCCAT
CGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTACTTGACATGGAGAAGTTCTG
CACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGCATTGT
AATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCACCACGTATTTGTGCAGATGAATCT
GGCTTCTTAGATCACTGCAGAAAAGGTTAAAGGCAAGGGGGAAGAGGTCTTGAGAG

AI911549

TTTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTACAGTTTTCAATTAAAGCT
GTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTAG
TTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTAC
TAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAA
TTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAG
CAGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGA

AGTTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACA

BF194822

TTCTCTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTACAGTTTTCAATTAAA
GCTGTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGT
TAGTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAAT
TACTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTG
TAATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAG
CAGCAGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGG
AGAAGTTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGG

AI034244

TTTTTTTTTTTTTTTTACAACCTTGAAAGCTGTTTTATACATAGAAATCAATTACAGTTT
TAATTGAAAACTATAACCATTTTGATAATGCAACAATAAAGCATCTTCAGCCAAACATCT
AGTCTTCCATAGACCATGCATTGCAGTGTACCCAGAACTGTTTAGCTAATATTCTATGTT
TAATTAATGAATACTAACTCTAAGAACCCCTCACTGATTCACTCAATAGCATCTTAAGTG
AAAAACCTTCTATTACATGCAAAAAATCATTGTTTTTAAGATAACAAAAGTAGGGAATAA
ACAAGCTGAACCCACTTTTACTGGACCAAATGATCTATTATATGTGTAACCACTTGTATG
ATTTGGATTTGCATAAGACCTTCCCTCTACAAACTAGATTCATATCTTGATTCT

AI033911

TTTTTTTTTTTTTTTTACAACTGCAAAGCTGTTTTATACATAGAAATCAATTACAGTTTT
AATTGAAAACTATAACCATTTTGATAATGCAACAATAAAGCATCTTCAGCCAAACATCTA
GTCTTCCATAGACCATGCATTGCAGTGTACCCAGAACTGTTTAGCTAATATTCTATGTTT
AATTAATGAATACTAACTCTAAGAACCCCTCACTGATTCACTCAATAGCATCTTAAGTGA
AAAACCTTCTATTACATGCAAAAAATCATTGTTTTTAAGATAACAAAAGTAGGGAATAAA
CAAGCTGAACCCACTTTTACTGGACCAAATGATCTATTATATGTGTAACCACTTGTATGA
TTTGGTATTTGCATAAGACCTTCCCTCTACAAACTAGATTCATATCTTGATTCT

BF064177

TTTTTTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAAGGCT
GTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATTGGTGG
GATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCAGCCAT
CGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTACTTGACATGGAGAAGTTCTG
CACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGCATTGT
AATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCACCACGTATTTGTGCAGATGAATCT
GGCTTCTTAGATCACTGCAGAAAAGGTTAAAGGCAAGGGGGAAGAGGTCTTGAGAGTTCT
CACTGGGACTGCCCTCGCTCTTGCCACAGGTACCATCGCACACACTGTTGACGTCATTGG
AAAG

AA847767

GGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAACTGTA
ATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTAGTTA
ATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAA
GGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATTG
GTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCAG
CCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAAGT
TCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTA

AI538624

TTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTACAGTTTTCAATTAAAGCTG
TAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTAGT
TAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACT
AAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAAT
TGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGC
AGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAA
GTTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTAC

AI913613

TTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAACTG
TAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTAGT
TAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACT
AAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTNTTTCCCTGTAAT
TGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGC
AGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAA
GTTCTGCACAGAAAGCAGTGGCATCCTTCATG

AI942234

GTTTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAAC
TGTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTA
GTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTA
CTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTA
ATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCA
GCAGCCATCGTGGCAGCTTGGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGAAG
AAGTTCTGCACAGAAAGCAGTGGCAT

AI580483

GTTTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAAC
TGTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTA
GTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTA

CTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTA
ATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCA
GCAGCCATCGTGGCAGGCTTGGATCTTTTTCCTGCTGACACCTGCTGCTTGACATTGGAA
AGTTCTGCACAGAAAGCAGTGGCATC

AI831909

TTTTGGCTGATGATGCTTTATTGTTGCATTATCAAAATGGTTACAGTTTTCAATTAAAGC
TGTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTA
GTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTA
CTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTA
ATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCA
GCAGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAG
AAGTTCTGCACAGAAAGCAGTGGCAT

AI672344

GGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTACAGTTTTCAATTAAAGCTGTA
ATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTAGTTA
ATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAA
GGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATTG
GTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCAG
CCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAAGT
TCTGCACAGAAAG

AW025192

GATTGGCTGTTTTATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAA
CTGTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTT
AGTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATT
ACTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGT
TATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGC
AGCAGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGA
GAAGTTCTGCACAAAAAGCAGTGGCATCCTTCATGAGGTGGTA

AA677205

GCAATATTTTAAAATTTTATTGGTATAAAATGTTTTAATTACTAAGGCTGTTTGTAGGCT
GCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATTGGTGGCATAGGAAGCC
TTTAAGGTCTCTTGCTTCTCATGGTGTGGGCTACAAGGAGCAGCAGCCATCGTGGCAGGC
TTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAAGTTCTGCACAGAAAGC
AGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGCATTGTAATCGTCTTT
TGTATCAATCTCTCTAAAGTAGACCACCACGTATTTGTGCAGATGAATCTGGCTTCTTAG
ATCACTGCAGAAAAGGTTAAAGGCAAGGGGGAAGAGGTCTTGAGAGTTCTCACTGGGACT

GCCCTCGCTCTTGCCACAGGTACCATCGCACACACTG

AA721647

TTTTTTTTTTACAACTTCAAAGCTGTTTTATACATAGAAATCAATTACAGTTTTAATTGA
AAACTATAACCATTTTGATAATGCAACAATAAAGCATCTTCAGCCAAACATCTAGTCTTC
CATAGACCATGCATTGCAGTGTACCCAGAACTGTTTAGCTAATATTCTATGTTTAATTAA
TGAATACTAACTCTAAGAACCCCTCACTGATTCACTCAATAGCATCTTAAGTGAAAAACC
TTCTATTACATGCAAAAAATCATTGTTTTTAAGATAACAAAAGTAGGGAATAAACAAGCT
GAACCCACTTTTACTGGACCAAATGATCTATTATATGTGTAACCACTTGTATGATTTGGT
ATTTG

BF115018

GTTTCGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAAC
TGTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTA
GTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTA
CTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTAAGGCCCATCACACGTTTTTTCCCTGTA
ATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTNTCATGGGTGGGCTACAAGGAGCA
GCAGCCATCGTGGCAGGCTTGNGATCTTTTTCCTGCTGGCCCCTGCTGCTTGACAT

W61238

NAAAGCACTGGCTGAAGGAAGCCAAGAGGATCACTGCTGCTCCTTTTTTCTAGAGGAAAT
GTTTGTCTACGTGGTAAGATATGACCTAGCCCTTTTAGGTAAGCGAACTGGTATGTTAGT
AACGTGTACAAAGTTTAGGTTCAGACCCCGGGAGTCTTGGGCACGTGGGTCTCGGGTCAC
TGGTTTTGACTTTAGGGCTTTGTTACAGATGTGTGACCAAGGGGAAAATGTGCATGACAA
CACTAGAGGTATGGGCGACACGANAACGAACGGGAAGTTTTGGCTGAAGTAGGAGTCTTG
GTGAGATTTTGCTCTGATGCATGGTGTGAACTTTCTGAGCCTCTTGTTTTTCCTCAAGCT
GACTCCATATTTTCCTACTTGTGGCAGCGACTGCATCCGACATAAAGGAACAG

W61239

TAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAAGGCTGTTTGTAGG
CTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATTGGTGGGATAGGAAG
CCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCAGCCATCGTGGCAGG
CTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAAGTTCTGCACAGAAAG
CAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGCATTGTAATCGTCTT
TTGTATCAATCTCTCTAAAGTAGACCACCACGTATTTGTGCAGATGAATCTGGCTTCTTA
GATCACTGCAGAAAAGGTTAAAGGCAAGGGGGGA

AI032064

AGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAAGGCTGTTTGTAGGC

TGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATTGGTGGCATAGGAAGC
CTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCAGCCATCGTGGCAGGC
TTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAAGTTCTGCACAGAAAGC
AGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGCATTGTAATCGTCTTT
TGTATCAATCTCTCTAAAGTAGACCACCACGTATTTGTGCAGATGAATCTGGCTTCTTAG
ATCACTGCAGAAAAGGTTAAAGGCAAGGGGGAAGAGGTCTTGAGAGTTCTCACTGGGACT
GCCCTCGCTCTTGCCAC

AW236941

TTTTTTTTTTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAAGGC
TGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATTGGTG
GGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCAGCCA
TCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAAGTTCT
GCACAAAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGCATTG
TAATCGTCTTTTGTATCAATC

AW236941

TTTTTTTTTTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAAGGC
TGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATTGGTG
GGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCAGCCA
TCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAAGTTCT
GCACAAAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGCATTG
TAATCGTCTTTTGTATCAATC

BG057174

TTTTATACATAGAAATCAATTACAGCTTTAATTGAAAACTATAACCATTTTGATAATGCA
ACAATAAAGCATCTTCAGCCAAACATCTAGTCTTCCATAGACCATGCATTGCAGTGTACC
CAGAACTGTTTAGCTAATATTCTATGTTTAATTAATGAATACTAACTCTAAGAACCCCTC
ACTGATTCACTCAATAGCATCTTAAGTGAAAAACCTTCTATTACATGCAAAAAATCATTG
TTTTTAAGATAACAAAAGTAGGGAATAAACAAGCTGAACCCACTTTTACTGGACCAAATG
ATCTATTATATGTG

AW058532

GGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAACTGTA
ATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTAGTTA
ATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAA
GGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCCTGTATGG
GTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCT

T98360

TNAGGAANGAGAAGAAGCGAGATNNANNTNNAGAAATANGTGGTGGCNTANTTTAGAGAG
ATTGATNCAAAAGCNGATTNCAATNNNCTCAGTGNCTNCCCAAGTNCCNCCTCATGAAGG
ATNCACTNCTTTCTGTGCAGACTNNNCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGAN
CACAAGCTCCNCGATGGCTGCTGCTCCTTGTAGCCCNCCATGAGAAGCAAGAGNCTTAAA
GGCTTCCTATCCCACCAATTACAGGGAAAAACGTGTGATGACCTGAGCTTACTATGCAGC
CTACAANCAGCCTTAGTAATTAAACCNTTTATT

T98361

NANNATGAAGATGCTTTATTGTTGCATTATCAAAATGGTTACAGTTTTCAATTAAAGCTG
TAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTAGT
TAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGNATAAAATGTTTTAATTACT
AAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTNCCCTGTAAT
TGGGTGGGGATAGGGAAGCCCTTTAAGGGTCTCTTGCTTCTCATGGGGTGGGGCCTACNA
AGGGAGCAGCCAGCCCATCGTGGCCAGGGCCTTGTGGANCCTTTTTCCCTGCCTGGACAC
CCTGCCTGCCTTGGACCATGGGGAGGAAGGTTCTGGCACCAGGAAAGCCAGGTGGCCCAT
CCCTTCCATGAGGGTGGGGTACTTNGGGGGGCCAGGACCACTGAGGNGCCATTGGTAATC
CGTCCTTTTNGTATCCAATCCCCTCCTAAGGTAGGNCCCCCC

AI470845

TTTTGTGGGTTCAGCTTGTTTATTCCCTACTTTTGTTATCTTAAAAACAATGATTTTTTG
CATGTAATAGAAGGTTTTTCACTTAAGATGCTATTGAGTGAATCAGTGAGGGGTTCTTAG
AGTTAGTATTCATTAATTAAACATAGAATATTAGCTAAACAGTTCTGGGTACACTGCAAT
GCATGGTCTATGGAAGACTAGATGTTTGGCTGAAGATGCTTTTATTGTTGCATTATCAAN
ATGGTTTATAGTTTTCAATTAAAACTGTAATTGATTT

AI497731

GGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAACTGTA
ATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTAGTTA
ATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAA
GGCTGTTTGTAGGCTGCATAGTAAGCTTAANGATCATACNCACGTTTTTCCCTGAATTTG
GTGGGATAANGAAGCCTTTAAAGGT

T96629

TTGAAAATTTTATTGGNATAAAATGTTTTAATTACTAAGGCTGTTTGTAGGCTGCATAGT
AAGCTTCAGGANCATCACACGTTTTTTCCCTGTAATTGGTGGCATAGGAAGCCTTTAAGG
TCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCAGCCATCGTGGCAGGCTTGTGATC
TTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAAGTTCTGCACAGAAAGCAGTGGCAT
CCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGCATTGTAATCGTCTTTTGTATCAA
TCTCTCTAAAGTAGACCACCACCGTNTTTGTGCAGATGGANTCTGGCTTC

T96740

AGGCACTATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGC
TTCAGTGGTGATTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATA
TTTTCCTACTTGTGGCAGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCACA
AACAGGCGTCCCTTTCCCTCTGGATAACAACAAAAGCAAGCCGGGANGGNCTGNCCTCTC
CTCCTGCTGTCTCTGCTGGTGGCCACATGGGTGCTGGTGGCAGGGATCTATCTAATGTGG
AGGCACGAAAGGATCAAGAAGACTTCCTTTTCTAACCACCACATTACTGCCCCCCATTTA
AGGTTCTTGTGGTTTTACCCATCTGGAAATATGTTTTCCCTTCACACATTTGTTTATTTC
ATTGATTTNTTTCAAAACCTTGGCAGGAGTTT

H25975

GGGTCCAGTGCAGTGGCTTGCNTGCAGAAAGAAGGCAGCAGACAAAGTCGTCTTCCTTCT
TTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCCAG
TGAGAACTCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAAG
CCAGATTCATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACGA
TTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTTCTG
TGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATTCACAAGCCTGCC
ACGATGGCTGCTTGCTTCCTTTGTAGCCCACCCATGAGGAAGNCAAGAGACCTTNAAAGG
GTTCCTTTTCCCATCANTTTACAGGGGANAAAACGTGTGATGATC

H25941

TTTTGTTTGGCTNATNTNNTTCTTATTGTTGCATTATCAAAATGGTTATAGTTTTCAATT
AAAACTGTAATTGATTNCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAAT
CGTTAGTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAANGTTTT
AATTACTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTCCC
CTGTAATTGGTGGGATAGGAAGCCTTTAAGGTCTCTNGCTTCTCATGGGTGGGCTACAAG
GAGCAGCAGCCATCGTGGCAGGCTTGTGANCTTTTNCCTGCTGACACCTGCTGCTTGACA
TGGGAGAAGTTCTGCACAGAAAGGCAGTGGGCATCCTTCATGAGGTGGGTACTTGGGGGN
CAGACACTGAGGAGCATTGT

BE539514

ACTCAAAAGAAGGCAGCAGACAAAGTCGTCTTCCTTCTTTCCAATGACGTCAACAGTGTG
TGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTTC
CCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAAGCCAGATTCATCTGCACAAATAC
GTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACGATTACAGTGCTCTCAGTGTCTGC
CCCAAGTACCACCTCATGAAGGATGCCACTGCTTTCTGTGCAGAACTTCTCCATGTCAAG
CAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACGATGGCCGCTGCTCCTTGTAG
CCCACCCATGAGAAGCAAGAGACCTTAAAGGCTTCCTATCCCACCAATTACAGGGAAAAA
ACGTGTGATGATCCTGAAGCTTACTATGCAGCCTACAAACAGCCTTAGTAATTAAAACAT
TTTATACCAATAAAATTTTCAAATATGCTAACTAATGTAGCATTAACTAACGATTGGAAA
CTACATTTACAACTTCAAAGCTGTTTTATACATAGAAATCAATTACAGCTTTAATTGAAA

ACTGTAACCATTTTGATAATGCAACAATAAAGCATCTTCAG

BX282554

GTCCAGTGCAGTGGCTTGCCACTCAAAAGAAGGCAGCAGACAAAGTCGTCTTCCTTCTTT
CCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCCAGTG
AGAACTCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAAGCC
AGATTCATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACGATT
ACAGTGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTTCTGTG
CAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCCGCTGCTCCTTGTAGCCCACCCATGAGAAGCAAGAGACCTTAAAGGCTTCCTATC
CCACCAATTACAGGGGAAAAAACGTGTGATGATCCTGAAGCTTACTAT

R74038

TATTGTTGCATTATCAAAATGGTTATAGTTTTCAATTAAAACTGTAATTGATTTCTATGT
ATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTAGTTAATGCTACATTAGTT
AGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTACTAAGGCTGTTTGTAGGC
TGCATAGTAAGCTTCAGGATCATCACACGTTTTTNCCCTGTAATTGGGTGGGGATAGGGA
AGCCTTTAAGGTCTCTTGCTTCTCATGGGGTGGGGCTACAAGGGAGGCAGGCAGCCATCG
TGGGCAGGGCTTGTGATCTTTTTCCCTGCTGACACCTGCTGCTTGACATGGGGGGAAGGT
TCTGGCACAGAAAGCAGTGGGCATCCTTCATGAGGGTGGTACTTGGGGGGCAGACACTGA
GGAGGCNTTGTAAATCGNCTTTTTNGTATCCAANCTCTNCTAAAGTAGGGNCCACCNCGT
TTTTTNTTGCAGGTGGATNCGGGGCTN

R74129

GGGTCCAGTGCAGTGGCTTGCNTNCAAAAGAAGGCAGCAGACAAAGTCGTCTTCCTTCTT
TCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCCAGT
GAGAACTCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAAGC
CAGATTCATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACGAT
TACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTTCTGT
GCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCAC
GATNGCTGCTGCTCCTTGTAGNCCACCCATGAGAAGCAAGTGACCTTTAAAGGNTTTCCT
ATTNCCACCNATTTACAGGG

BG433769

GACTAGATGTTTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAGTTTTCA
ATTAAAACTGTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCC
AATCGTTAGTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGT
TTTAATTACTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTT
TCCCTGTAATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTAC
AAGGAGCAGCAGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTG

ACATGGAGAAGTTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAG
ACACTGAGAGCATTGTAATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCACCACGTAT
TTGTGCAGATGAATCTGGCTTCTTAGATCACTGCAGAAAAGGTTAAAGGCAAGGGGGAAG
AGGTCTTGAGAGTTCTCACTGGGACTGCCCTCGCTCTTGCCACAGGTACCATCGCACACA
CTGTTGACGTCATTGGAAAAGAAGGAAGAC

BG530489

GAGTTCTCACTGGGACTGCCCTCGCTCTTGCCACAGGTACCATCGCACACACTGTTGACG
TCATTGGAAAGAAGGAAGACGACCTTGTCTGCTACCTTCTTTTGAGTGGCAAGCCACTGC
ACTGGACCCATCTCTGCTATTTTCTTTTTCTGCCACTTTTCAAGGATGACCTCACTTCTG
CAATGGTTTTGAAGAAATTCAGTGAAGTAACAAATTGTGTGATGGAAACATATTTCAGAT
GGGTAAACCACAAGAACCTTAATGGGGGGCAGTAGTGTGGTGGTAGAAAAGGAAGTCTTC
TTGATCCTTTCTGTGAGAGGAGAAAAGCATTTGTTATCTGTGAATAGCAAACAGCAGGCT
TTCACTCTGTAAACCATCCCTGACAAATGATCCCTTGCTAGAGAATGTCAGCTGAGCACC
AAGGGCCTTGTTAGTGACAGCAAGGAAAAACATCCTGATGTTCCTTTTGAACACATCACC
TGAAACACACTGATGCTTAAACCTTAACTTTTTTTTTTTGGGGGACATAGTCTCACTCTG
TCGCCCAGGCTGGAGTGCGTGGGAGAGGACCTCGGAAAGACTGGCAAGCATCCGCATACA
AGGGAGTAACAGCACAATACTCCGTGAACTTCGGAGCCCTCCAAAGGAATACTCAAGGGC
GGGTAAAGGATGGCAAGGGTCGACGGAGAGCCCACGAGGAGAGCGGAAGGTAGAGAGGAG
ACAAGCATAAGACGCGAGAGGAACTCCAAGGCGGGGCCAAAGAGAGAAACCACGGTCACC
AACAGAAG

AA007528

AGAAGCCAGATTCATCTGCACAAATACGTGGTGNTCTACTTTAGAGAGATTGATACAAAA
GACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCT
TTCTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCC
TGCCACGATGGCTGCTGCTCCTTGTAGCCCACCCATGAGAAGCAAGAGACCTTAAAGGCT
TCCTATCCCACCAATTACAGGGNAAAAACNGTAGTGATNATCCCTGACAGCTTACTATGC
CAGCCNT

AA007529

TTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATCGGTTACAGTTTTCAATTAAAGCT
GTAATTNGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGTTTCCAATCGTTA
GTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATAAAATGTTTTAATTA
CTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTA
ATTGGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATTGGGTGGGCTACAAGGAG
CAGCAGCCATCCGTNGGCAAGGCTTTGTGGATNCT

BI260259

GGAAGAGAAAGATCGTCCAGAGGTTCCATCGCACACACTGTATGACGTCATTGGAAATGA

AGGAAGACGACTTTGTCTGCTGGCTTCTTGTGAGTGGCAAGCCACTGCAGTGGACCCATC
TCTGCTATTTTCTTTATTCTGCCACTTTTCAAGGATGACCTCACTTCTGCAATGGTTTTG
AAGAAAGTTCAGTGAAGTAACAAATTGTGTGATGGAAACATATTTCAGATGGGTAAACCA
CAAGAACCTTAATGGGGGGCAGTAGTGTGGTGGTAGAAAAGGAAGTCTTCTTGATCCTTT
CTGTGAGAGGAGAAAAGCATTAGTTATCTGTGAACAGCAAACAGCAGGCATTTCACATCT
GTAAACCATCCCTGACAAATGATCCCTTGCTAGAGAATGTCAGCTGAGCACCAAGGGGCC
TTGTTAGTGACAGCAAGGACAAAACATCCTGATGTTCCTTTTGAACACATCAGCTGAAAC
ACACTGATGCTCTAAACCGTTAACTATTTATTAATGGGGGAACATAGGTCTCAACTCATG
TACGACCAGGCTGGAGTGCAGTGGGGTTGAACATCGACAGACATAGCAAACCACCGATCA
CTAGGGAAACAACGCACAGAACTCCAGACTTAAAACACC

AA287951

ATTCGGCACCTGGGGGGCAGACACTGAGAGCATTGTAATCGTCTTTTGTATCAATCTCTC
TAAAGTAGACCACCACGTATTTGTGCAGATGAATCTGGCTTCTTAGATCACTGCAGAAAA
GGTTAAAGGCAAGGGGGAAGAGGTCTTGAGAGTTCTCACTGGGACTGCCCTCGCTCTTGC
CACAGGTACCATCGCACACACTGTTGACGTCATTGGAAAGAAGGAAGACGACTTTGTCTG
CTGCCTTCTTTTGAGTGGCAAGCCACTGCACTGGACCCATCTCTGCTATTTTCTTTTTCT
GCCACTTTTCAAGGATGACCTCACTTCTGCAATGGTTTTGAAGAAATTCAGTGAAGTAAC
AAATNTGTGTGATGGAAACATATTTCAGATGGGTAAACCACAAGAACCTTAATGGGGGGC
AGTAGTGTGGTGGTAGAAAAGGAAGTCTTCTTGATCCTTTCTGTGAGAGGAGAAAGC

AA287911

TTTTGATGGTCCACTTCCATTTAATGAATTAGTAAATATCTTTTCTCATGATTTTAATTA
CATTTTTTTCTCTAGCTTACTTTATTATAATACAGCACATAATACACCTAACATGCAAAA
TATGTGTTAATTGGCTGTTTATGTTATTGGTAAGACTTCCAGTCAACAGTAGGCTATTAG
AAGTTAAGTTGTGGGAAAATCAAAGGTTATAGGAGATTTTCAACTGCATGCAGGGCCGGT
GCCCTCCCCACTGTGTTGTTCAAGGGTCAGCTGTACTCTCTAAGGGCTTTGCTAACTTCA
AAACATGGAGTATTTGAATACAGAAACCAGAGCATTTACATACTCAGCTCAAGGCAGAGC
TATTAAAAAAACTCCTCTTCTCCATATGTAGGAAAGGAAATACAAATGCATCCTTTGAGT
CATTTGTGATGT

T97852

AACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCTGGATAACAACAAAAGCAA
GCCGGGANGNCTGNCGCTCTCCTCCTGCTGTCTCTGCTGGTGGCCACATGGGTGCTGGTN
GCAGGGATCTATCTAATGTNGAGGCACGAAAGGGATCAAGAGGACTTCCTTTTCTACCAC
CACACTACTGCCCCCCATTAAGGTTCTTGTNGGTTTACCCATCTGGAAATATGTTTCCAT
CACACAATTTGTTACTTCACTGGAATTTCTTCAAAACCATTGGCAGGANGTGAGGGTCAT
CCTTGGAAAAGTGGGC

T97745

CCTCACTTCTGCAATGGTTTTGAAGAAATTCAGTGAAGTAACAAATTGTGTGATGGAAAC
ATATTTCAGATGGGTAAACCACAAGAACCTTAATGGGGGGCAGTAGTGTGGTGGTAGAAA
AGGAAGTCTTCTTGATCCTTTCGTGCCTCCACATTAGATAGATCCCTGCCACCAGCACCC
ATGTGGCCACCAGCAGAGACAGCAGGAGGAGAGGCAGCCAGCCTCCCGGCTTTGCTTTTG
TTGTTATCCAGAGGGGAAAGGGGACGCCTGTTTNTGGGGCAGAGCACAACTGTTTCCCTC
GTGCCCGAATTCTTTGGGCCTTCGAGGGGCCAAATTTCCCTATTAGGTGAGGTCGTATTT
TAAATTTCGGTAATTCATGGTCATAGGCTTGTTTTTCCCCG

N40294

GTTTCAACACAATTTTGGATCAGCTGCCTGTTTGCAAAAACATAATATATTTCTGTTAAA
CAGTTCTTCACCTAACAGCATATTGCTCTTATAACTGGTAGAGCTGTTTCAAAGGAAGTT
GGTTTCTGGTCCAAGTTTTGACCTAAACCATGTCCATCTTCTATTACCAGCACTTACAAG
CACTGTGAAAACTGATCATGACAAATAAGTAAAATTTGCTACATTAAACATATTGCCTCA
GCCATTACTAAGCGTCCACTTGTAAAGCTGGACACAGTTTTTACTTTATGCTTCATTTTG
ATTTTTTATCCGTAAGACATAAATTAGAAGGCATGAGGTGGCCCTTTAAGGATAATCTGC
AAATATACACATTTTAAATAGTCATCCATCTGGAAATCGNTCCACCATTCCAGGGGAAGG
ATTCCAGGTATTGGTGCTGTGGTGGAAATAAAGCATTCCCCNGGGAAAAAAACCATTTTA
TGNCTAAATAATTACCACCATTAACCTCNTGGGGTT

AA809841

GAATACTAACTCTAAGAACCCCTCACTGATTCACTCAATAGCATCTTAAGTGAAAAACCT
TCTATTACATGCAAAAAATCATTGTTTTTAAGATAACAAAAGTAGGGAATAAACAAGCTG
AACCCACTTTTACTGGACCAAATGATCTATTATATGTGTAACCACTTGTATGATTTGGGA
TTTGCAT

AA832389

TTTTTTACAACTTCAAAGCTGTTTTATACATAGAAATCAATTACAGTTTTAATTGAAAAC
TATAACCATTTTGATAATGCAACAATAAAGCATCTTCAGCCAAACATCTAGTCTTCCATA
GACCATGCATTGCAGTGTACCCAGAACTGTTTAGCT

H14692

CTGAGTGTGATGGTGTAAGCCTGTGGTCCCAGCTACTAGGGAGGCTGAGATGGGATTACA
GGTGTGAGCCACGGCGCCTGGCCTAAAAGCATCTTTTTCTTTAACGCAGAGGTTATGTTG
TATTATTAGCATAAATGTTTTTTTCTGGGAATGCTTATTTCACACAGCACAATACTGAAT
CTTCTCTGGAATGTGGATCGATTTCAGATGGATGACTATTAAAATGTGTATATTTGCAGA
TTATCCTTAAAGGGCCACCTCATGCCTTCTAATTTATGTCTTACGGATAAAAAATCAAAA
TGAAGCATAAAGTAAAAACTGTGTCCAGCTTTACAAGTGGACGCTTAGTAATGGCTGAGG
CAATATGTTTAATGTAGCCAAATTTTACTTATTTGTCCATGATCCAGTTTTTCACAGTGC
TTGTTAAGTGCTGGTAATTAGGAAGGTGGGACATGGGTTAGGTCAAAACTTGGGACCNGA
AACCAACTTGN

AA732635

TTTTTTTTTTACAACTTCAAAGCTGTTTTATACATAGAAATCAATTACAGTTTTAATTGA
AAACTATAACCATTTTGATAATGCAACAATAAAGCATCTTCAGCCAAACATCTAGTCTTC
CATAGACCATGCATTGCATTGTACCCAGAACTGTTTAGCTAATATTCTATGTTTAATTAA
TGAATACTAACTCTAAGAACCCCTCACTGATTCACTCAATAGCATCTTAAGTGAAAAACC
TTCTATTACATGCAAAAAATCATTGGTTTT

AA928257

TTTTCTGAGTAAGAACAGGCTTTATTTGTAAAACCACTCGTGACTCTTTACAAAGCAGGA
TACACAGAAGGGAAAAAAATACACAGTGCAAAATGGATGTTCTGAGTGCCACAAGGATCT
GCTGAAAAAAGCCAAAGATGTAAGATGGCTGGGTATATATGAGAATGAATATTTCACTAT
ATTCTGATTCAATTACCAGTCTCAGTGGCCCAGGATGAGCTTTTGGTGTGGTCACATGGC
CAACATTTGGATAACAAATGAGGAATAATGGTACCGCCTCACTAGTGCCTGAGAACAGCA
TGTTCTGGAAAATGTCTCTGGAGTTAGAGATGTGTTAGCTTTTTCATTACAGATGGAGAA
ATACAATGTTTACACAACAGTCCAGGGGTGGGGTCAAAAGTTGGAAGGTGTCATTAGACG
CAGCCAAATAAAGTGAAGACAACCCAGGTGACTGGCAGCCCTGACTTGTGCGTGGGCG

AI184427

TTTCTGAGTAAGAACAGGCTTTATTTGTAAAACCACTCGTGACTCTTTACAAAGCAGGAT
ACACAGAAGGGAAAAAAATACACAGTGCAAAATGGATGTTCTGAGTGCCACAAGGATCTG
CTGAAAAAAGCCAAAGATGTAAGATGGCTGGGTATATATGAGAATGAATATTTCACTATA
TTCTGATTCAATTACCAGTCTCAGTGGCCCAGGATGAGCTTTGGTGGTGGTCACATGGCC
AACATTTGGATAACAAATGAGGA

AI298577

GAGATGGAGGTCTCGCTTTGTGACGTAGCCTGGTCTTGAGCGATCCTTTTGCCTTGGCCT
TGCCAAAGTGCTGGGATTGGAGGCATGAGCCACTGCACCCACCCCTGTTTTTTTTTTAAG
TAAACCATTATAATAACTCATTTATAAAAAGGTTACTTCAAGAGGGCTTTCAACTTAAGA
ATTATTTTCATTTTGAACATGAAAAGTTAAATAGTAACTAAGAAACTGAGAACTCTGACA
GTGACCTCTAATAGGTAACTTTAGGCAAAAGTAGACAAGTTTGTGGGTATTTTGTTGTTC
ATGTTAAAAGGCACCTGTACAAGAATCAAGATATGAATCTAGTTTGTAGAGGGAAGGTCT
TATGCAAATACCAAATCATACAAGTGGT

AI692717

AGAGATGTTGGTCTCGCTTTGTGACGTAGCCTGGGCTTGAGCGATCCTTTTGCCTTGGCC
TTGCCAAAGTGCTGGGATTGGAGGCATGAGCCACTGCACCCACCCCTGTTTTTTTTTTAA
GTAAACCATTATAATAACTCATTTATAAAAAGGTTACTTCAAGAGGGCTTTCAACTTAAG
AATTATTTTCATTTTGAACATGAAAAGTTAAATAGTAACTAAGAAACTGAGAACTCTGAC

AGTGACCTCTAATAGGTAACTTTAGGCAAAAGTAGACAAGTTTGTGGGTATTTTGTTGTT
CATGTTAAAAGGCACCTGTACAAGAATCAAGATATGAATCTAGTTTGTAGAGGGAAGGTC
TTATGCAAATACCAAATCATACAAGTGGTTACACATATAATAGATCATTTGGTCCAGTAA
AAGTGGGTTCAGCTTGTTTATTCCCTACTT

AA910922

GAGATGGAGGTCTCGCTTTGTGACGTAGCCTGGTCTTGAGCGATCCTTTTGCCTTGGCTT
GCAAAGTGCTGGGATTGGAGGCATGAGCACTGCACCCACCCCTGTTTTTTTTTTTAAGTA
AACCATTATAATAACTCATTTATAAAAAGGTTACTTCAAGAG

H90761

TTCACTCAATAGCATCTTAAGTGAAAAACCTTCTATTACATGCAAAAAATCATTGTTTTT
AAGATAACAAAAGTAGGGAATAAACAAGCTGAACCCACTTTTACTGGACCAAATGANCTA
TTATATGTATAACCACTTGTATGATTTGGTATTTGCATAAGACCTTCCCTCTACAAACTA
GATTCATATCTTGATTCTTGTACAGGTGCCTTTTTAATATTCTGTGATGAAATCGTTCAC
AGTCAGAGTACATGTCTGCTGCATATGGGAAATAGGGACTGTTGTTCTGAGGGACAAGGC
ACTCAATTCAGCCGTAAAGGCTGACCCGGGCTACTTTTTTTCCANGGGAATACAATTTTT
TTACCTTGGAATAAAATNGGGCCCGACNGGAC

AI620122

TTTTTTTTTTTGAGTAAGAACAGGCTTTATTTGTAAAACCACTCGTGACTCTTTACAAAG
CAGGATACACAGAAGGGAAAAAAATACACAGTGCAAAATGGATGTTCTGAGTGCCACAAG
GATCTGCTGAAAAAAAGCCAAAGATGTAAGATGGCTGGGTATATATGAGAATGAATATTT
CACTATATTCTGATTCAATTACCAGTCTCAGTGGCCCAGGATGAGCTTTTGGTGTGGTCA
CATGGCCAACATTTGGATAACAAATGAGGAATAATGGTACCGCCTCACTAGTGCCTGAGA
ACAGCATGTTCTGGAAAATGTCTCTGGAGTTAGAGATGTGTTAGCTTTTTCATTACAGAT
GGAGAAATACAATGTTTACACAACAGTCCAGGGGTGGGGTCAAAAGTTGGAAGGTGTCAT
TAGACGCA

AI793318

AAATTTTTAACTTTTAATAGTTAAAATAGTTAACTATTGGTATGGTAGGAAATGATAAAG
TAGACTAGTATCTGTATACATTTTCTGCATTTATGACATACCTTTTTCTTCATTTTTTTC
AATATTTTAATTGAAAAGTTCATCCGAGTTTCATCTAAGTTTTTTCAAAGTGATACAAAT
CTCCAAAAAATTTTCCAATATATGTATTGAAAAAATCCAGGTGTAAGTGGCTCTGCGCAG
TCCAAACCTGTGTTGTTCAAGGGTCAACTGTGTATGAATCCAAGCGAAAGCTTTTCTTAA
CACCTCATAAGAACTATTTTTTAAAAAACAGGAACTAGCATAGAGTAACCATCACAGGTA
AAGTGTAATTTGTTATCAGCCATCTTTTGCCCATTTCAGTACTGGTAGAAGGCTCAATGG
TAAAAATAAA

AA962325

TTTTTTTTTTTTTTTTTTTTTTTCTGACTGTCCCGTTTTTATTTTTACCATTGAGCCTT
CTACCAGTACTGAAATGGGCAAAAGATGGCTGATAACAAATTACACTTTACCTGTGATGG
TTACTCTATGCTAGTTCCTGTTTTTTAAAAAATAGTTCTTATGAGGTGTTAAGAAAAGCT
TTCGCTTGGATTCATACACAGTTGACCCTTGAACAACACAGGTTTGGACTGCGCAGACCA
CTTACACCTGGATTTTTTCAATACATATATTGGAAAATTTTTTGGGGATTTGTATCACTT
TGAAAAAACTTAGATGAAACTCGGATGGACTTTTCCATTAAAATATTGGAAAAAATGAAG
AAAAAGGT

AI733290

TTTTTTTTTTTTTTTTTTTTTTTCTGACTGGCCCGTTTTTATTTTTACCATTGAGCCTT
CTACCAGTACTGAAATGGGCAAAAGATGGCTGATAACAAATTACACTTTACCTGGGATGG
TTACTCTATGCTAGTTCCTGTTTTTTAAAAAATAGTTCTTATGAGGGGTTAAAAAAAGCT
TTCGCTTGGATTCATACACAGTTGACCCTTGAACAACACAGGTTTGGACTGCGCAGAGCC
ACTTACACCTGGATTTTTTCAATACATATATTGGAAAATTTTTTGGAGATTTGTATCACT
TTGAAAAAACTTAGATGAAACTCGGATGAACTTTTCAATTAAAATATTGAAAAAAATGAA
GAAAAAGGTATGTCATAAATGCAGAAAATGTATACAGATACTAGTCTACTTTATCATTTC
CTACCATACCAATAG

BQ226353

TAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCTGGATAACAGTAA
GTGCCCAGTAACTTCAACCAGATGATCAAAGTGGCTCACACACAGTCACTGCCCCCCACT
CAGTATGTGGAAGGGTTGTGTGTATGTGGGCAGTGCAAGGGGTCGCTGCCTGTGTACACT
GAACTGGGGTGCAGAGAAAGCCAACAGTGCTGTCCCAGAGAACCTAGAATCTGAGTAAGA
ACAGGCTTTATTTGTAAAACCACTCGTGACTCTTTACAAAGCAGGATACACAGAAGGGAA
AAAAATACACAGTGCAAAATGGATGTTCTGAGTGCCACAAGGATCTGCTGAAAAAAGCCA
AAGATGTAAGATGGCTGGGTATATATGAGAATGAATATTTCACTATATTCTGATTCAATT
ACCAGTCTCAGTGGCCCAGGATGAGCTTTTGGTGTGGTCACATGGCCAACATTTGGATAA
CAAATGAGGAATAATGGTACCGCCTCACTAGTGCCTGAGAACAGCATGTTCTGGAAAATG
TCTCTGGAGTTAGAGATGTGTTAGCTTTTTCATTACAGATGGAGAAATACAATGTTTACA
CAACAGTCCAGGGGTGGGGGTCAAAAGTTGGAAGGTGTCATTAGACGCAGCCAAATAAAG
TGAAGACCACCCAGGTGACTGGCAGCCCTGACTTGTGCGTGGGCGAAACCTTACAGATTC
CTGGGGCACTCTGTGCCTGAACTTACCTGGATGGTCTTTGTGAGGCGGGTGGGCACTTAT
CCTCCATNAATGGTCAGTCTAACAAGACCGGCCTGTAAAAATGGCATCTAATAGGGGCTA
TGGAATGGAAAACAGTTGGTACCCAGAAATAACTTTAATT

W04890

GACAGTCTGGGAGCCCAGAGCTCTGGGAGGAGTNGGGAAAATGCTGCTTCCTGCTGCTTG
CTTCTAGGCACCTGCTTCCGCCATCTCACTTACCATGGCTAGAGATGGGGGTGAGACTGG
GGAAGGACAAAAGCAGGGAACAGATAAGGGATGGAAATCAGAAGGGAATATAGAAAGAAC
TCTGGATATGCNAGAAATGCCGGTACCTGAGCATTTTGTATCAATGGGAGTACCCTCTGT
AACTGCTCAGTAGGTTACAAATGAAGAGTCCACCAGTATTAGAAACAATTTAAACTTGCC

AGTACCAACTGGGATGTGTGCCTTCAATTTGAAAATTGTATGTTTTATTTTTTAAATTTG
GTTAACAGCATTAATTTATAGAGTATTTGATGTCATTTATGGTTCCCGAGGTGTTTCCAA
CACAATTTTTGGGATCA

BM455231

CTTTTAATAGTTAAAATAGTTAACTATTGGTATGGTAGGAAATGATAAAGTAGACTAGTA
TCTGTATACATTTTCTGCATTTATGACATACCTTTTTCTTCATTTTTTTCAATATTTTAA
TTGAAAAGTTCATCCGAGTTTCATCTAAGTTTTTTCAAAGTGATACAAATCTCCAAAAAA
TTTTCCAATATATGTATTGAAAAAATCCAGGTGTAAGTGGCTCTGCGCAGTCCAAACCTG
TGTTGTTCAAGGGTCAACTGTGTATGAATCCAAGCGAAAGCTTTTCTTAACACCTCATAA
GAACTATTTTTTAAAAAACAGGAACTAGCATAGAGTAACCATCACAGGTAAAGTGTAATT
TGTTATCAGCCATCTTTTGCCCATTTCAGTACTGGTAGAAGGCTCAATGGTAAAAATAAA
AACGGGACAGTCAGAAGATCTGGAAGTCCTGACCCTGCTTTCACCTGGCATGTGTAATCC
AGTCATGCTCGTATCAGTCTCTGTAGGAGCACTTGAAGGTATTACATAAATGCTATCTAA
CTCTGGGAAACGCCAACATGTGATTGCCTCCAGAGGAATCTTCTTTAAAAAAAAATTCAA
AATGTTATTTCCTTACTAGGATGTCTTTAAAGAATTATAACCCTTACCGTGCCTCCACAT
TAGATAGATCCCTGCCACCAGCACCCATGTGGCCACCAGCAGAGACAGCAGGAGGAGAGG
CAGCCAGCCTCCCGGCTTGCTTTTGTCTGGAAAAAAACAAAGCTTATTCACCTTTGGAAA
AAAATCCACACTTATCTCTTAATTTAAAAACTAAGACTTGGTATACTTTATAGAGGGTTA
TTTATTTTTTATTATTTTTTAGTTTTGAGACAGAGTCTCGCTTTGTTGCCTANGCTGGAG
TGCAGTGGCGCAATCTCGGTTCACTGCAGCCTCCGTTCTCCGGGGTTCAAGGCATGCTGG
CTCAGCCTCCTGTATAGCTGGGGATTAAAGGCATGTGTTCACGCGGCCCAGCCCCTTTTG
TAAAAGATTTAGATCCCTTTTAAAACCATCAGTCAGGAGGCTCCTTTAAAAAGTCTGGCC
ATCTAATCTTTTTTCCCCCAAAAGGGG

BI492426

TTTTTTTTTTTCTTTTTTCTGAGTAAGAACAGGCTTTATTTGTAAAACCACTCGTGACTC
TTTACAAAGCAGGATACACAGAAGGGAAAAAAATACACAGTGCAAAATGGATGTTCTGAG
TGCCACAAGGATCTGCTGAAAAAAGCCAAAGATGTAAGATGGCTGGGTATATATGAGAAT
GAATATTTCACTATATTCTGATTCAATTACCAGTCTCAGTGGCCCAGGATGAGCTTTTGG
TGTGGTCACATGGCCAACATTTGGATAACAAATGAGGAATAATCTCGTGC

BG674622

AATTTATAGAGTATTGATGTCATTTATGTTTCTGAGGTGTTTCAACACAATTTTGGATCA
GCTGCCTGTTTGCAAAAACATAATATATTTCTGTTAAACAGTTCTTCACCTAACAGCATA
TTGCTCTTATAACTGGTAGAGCTGTTTCAAAGGAAGTTGGTTTCTGGTCCAAGTTTTGAC
CTAAACCATGTCCATCTTCTATTACCAGCACTTACAAGCACTGTGAAAACTGATCATGAC
AAATAAGTAAAATTTGCTACATTAAACATATTGCCTCAGCCATTACTAAGCGTCCACTTG
TAAAGCTGGACACAGTTTTTACTTTATGCTTCATTTTGATTTTTTATCCGTAAGACATAA
ATTAGAAGGCATGAGGTGGCCCTTTAAGGATAATCTGCAAATATACACATTTTAATAGTC
ATCCATCTGAAATCGATCCACATTCCAGAGAAGATTCAGTATTGTGCTGTGTGAAATAAG

CATTCCCAGAAAAAAAACATTTATGCTAATAATACAACATAACCTCTGCATTAAAGAAAA
AGATGCTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCCTGCACTTTGAGAGGCTG
AGGTGGGTGGATCATGAGGTCAGGAGATCAAGACCATCCTGGCTAACAGGGTGAAACCCC
GTCTCTACTGGGGATATAACAAAGTTAGCTGGGTGTGGTGGTGGGTGCTTGTGGTCCCAG
CTACTCAGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGGAAGGCAGAGGTTGTAGTGAC
GCGAGGTTCACGCCACTGCATTCCAGTCTGGG

BX111256

CAGGAAGNTAAGAACAGTCCTAAAATCTCTTTGGCTTCTTTGTCCTGATATGCACCGGCA
TTTTCACAGTAGGAACTAGGGTTTCTGTCCAGTTTTTTTGGTTCTTTAAGGAATTAATGT
TATTCTGGGTACAACTGCTTACATACATAGCACATATAGATGACATTTTTACAGGCCGTC
TTGTTAGACTGACATACATGGAGGATAGTGCCACCCGCCTCACAAGAACATCAGGTAAGC
TCAGGCACAGAGTGCCCAGGAATCTGTAAGGCTTCGCCCACGCACAAGTCAGGGCTGCCA
GTCACCTGGGTTGTCTTCACTTTATTTGGCTGCGTCTAATGACACCTTCCAACTTTTGAC
CCCACCCCTGGACTGTTGTGTAAACATTGTATTTCTCCATCTGTAATGAAAAAGCTAACA
CATCTCTAACTCCAGAGACATTTTCCAGAACATGCTGTTCTCAGGCACTAGTGAGGCGGT
ACCATTATTCCTCATTTGTTATCCAAATGTTGGCCATGTGACCACACCAAAAGCTCATCC
TGGGCCACTGAGACTGGTAATTGAATCAGAATATAGTGAAATATTCATTCTCATATATAC
CCAGCCATCTTACATCTTTGGCTTTTTTCAGCAGATCCTTGTGGCACTCAGAACATCCAT
TTTGCACTGTGTATTTTTT

BX117618

AAATTTTTAACTTTTAATAGTTAAAATAGTTAACTATTGGTATGGTAGGAAATGATAAAG
TAGACTAGTATCTGTATACATTTTCTGCATTTATGACATACCTTTTTCTTCATTTTTTTC
AATATTTTAATTGAAAAGTTCATCCGAGTTTCATCTAAGTTTTTTCAAAGTGATACAAAT
CTCCAAAAAATTTTCCAATATATGTATTGAAAAAATCCAGGTGTAAGTGGCTCTGCGCAG
TCCAAACCTGTGTTGTTCAAGGGTCAACTGTGTATGAATCCAAGCGAAAGCTTTTCTTAA
CACCTCATAAGAACTATTTTTTAAAAAACAGGAACTAGCATAGAGTAACCATCACAGGTA
AAGTGTAATTTGTTATCAGCCATCTTTTGCCCATTTCAGTACTGGTAGAAGGCTCAATGG
TAAAAATAAAAACGGGACAGTCAGAAAAA

AA682806

TCTGAGTAAGAACAGGCTTTATTTGTAAAACCACTCGTGACTCTTTACAAAGCAGGATAC
ACAGAAGGGAAAAAAATACACAGTGCAAAATGGATGTTCTGAGTGCCACAAGGATCTGCT
GAAAAAAGCCAAAGATGTAAGATGGCTGGGTATATATGAGAATGAATATTTCACTATATT
CTGATTCAATTACCAGTCTCAGTGGCCCAGGATGAGCTTTTGGTGTGGTCACATGGCCAA
CATTTGGATAACAAATGAGGAATAATGGTACCGCCTCACTAGTGCCTGAGAACAGCATGT
TCTGGAAAATGTCTCTGGAGTTAGAGATGTGTTAGCTTTTTCATTACAGATGGAGAAATA
CAATGTTTACACAACAGTCCAGGGGTGGGGTCAAAG

AI202376

CTGACTGTCCCGTTTTTATTTTTACCATTGAGCCTTCTACCAGTACTGAAATGGGCAAAA
GATGGCTGATAACAAATTACACTTTACCTGTGATGGTTACTCTATGCTAGTTCCTGTTTT
TTAAAAAATAGTTCTTATGAGGTGTTAAGAAAAGCTTTCGCTTGGATTCATACACAGTTG
ACCCTTGAACAACACAGGTTTGGACTGCGCAGAGCCACCCTCGTGCCGAATT

AI658949

CTGACTGTCCCGTTTTTATTTTTACCATTGAGCCTTCTACCAGTACTGAAATGGGCAAAA
GATGGCTGATAACAAATTACACTTTACCTGTGATGGTTACTCTATGCTAGTTCCTGTTTT
TTAAAAAATAGTTCTTATGAGGTGTTAAGAAAAGCTTTCGCTTGGATTCATACACAGTTG
ACCCT

BG403405

GGAAATGATAAAGTAGACTAGTATCTGTATACATTTTCTGCATTTATGACATACCTTTTT
CTTCATTTTTTTCAATATTTTAATTGAAAAGTTCATCCGAGTTTCATCTAAGTTTTTTCA
AAGTGATACAAATCTCCAAAAAATTTTCCAATATATGTATTGAAAAAATCCAGGTGTAAG
TGGCTCTGCGCAGTCCAAACCTGTGTTGTTCAAGGGTCAACTGTGTATGAATCCAAGCGA
AAGCTTTTCTTAACACCTCATAAGAACTATTTTTTAAAAAACAGGAACTAGCATAGAGTA
ACCATCACAGGTAAAGTGTAATTTGTTATCAGCCATCTTTGCCCATTTCAGTACTGGTAG
AAGGCTCAATGGTAAAAATAAAAACGGGACAGTCAGAAGATCTGGAAGTCCTGACCCTGC
TTTCACCTGGCATGTGTAATCCAGTCATGCTCGTATCAGTCTCTGTAGGAGCACTTGAAG
GTATTACATAAATGCTATCTAACTCTGGGAAACGCCAACATGTGATTGCCTCCAGAGGAA
TCTTCTTTAAAAAAAAATTCAAAATGTTATTTCCTTACTAGGATGTCTTTAAAGAATTAT
AACCCTTACCGTGCCTCCACATTAGATAGATCCCTGCAACAGACCCATGTGGCACCAGCA
GAGACAGCAGGAGGAGAGGCAGCAGCTCCCGGTTGTTTGTCTGGAAAAACAAAGGTTATC
ACTTTG

BE673417

CTGACTGTCCCGTTTTTATTTTTACCATTGAGCCTTCTACCAGTACTGAAATGGGCAAAA
GATGGCTGATAACAAATTACACTTTACCTGTGATGGTTACTCTATGCTAGTTCCTGTTTT
TTAAAAAATAGTTCTTATGAGGTGTTAAGAAAAGCTTTCGCTTGGATTCATACACAGTTG
ACCCT

AW021469

GCACGAGATTATTCCTCATTTGTTATCCAAATGTTGGCCATGTGACCACACCAAAAGCTC
ATCCTGGGCCACTGAGACTGGTAATTGAATCAGAATATAGTGAAATATTCATTCTCATAT
ATACCCAGCCATCTTACATCTTTGGCTTTTTTCAGCAGATCCTTGTGGCACTCAGAACAT
CCATTTTGCACTGTGTATTTTTTTCCCTTCTGTGTATCCTGCTTTGTAAAGAGTCACGAG
TGGTTTTACAAATAAAGCCTGTTCTTACTCAGAAAAAAAAAAAAAAAAAAA

CF455736

NNTTGAACAGGCGTGACGGTCCGGATTCCCGGGATGTTGTGCTCTGCCCACAAACAGGCG
TCCCTTTCCCTCTGGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGC
TGTCTCTGCTGGTGGCCACATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACG
AAAGGATCAAGAAGACTTCCTTTTCTACCACCACACTACTGCCCCCCATTAAGGTTCTTG
TGGTTTACCCATCTGAAATATGTTTCCATCACACAATTTGTTACTTCACTGAATTTCTTC
AAAACCATTGCAGAAGTGAGGTCATCCTTGAAAAGTGGCAGAAAAAGAAAATAGCAGAGA
TGGGTCCAGTGCAGTGGCTTGCCACTCAAAAGAAGGCAGCAGACAAAGTCGTCTTCCTTC
TTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCCA
GTGAGAACTCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAA
GCCAGATTCATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACG
ATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTTCT
GTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCC
ACGATGGCTGCTGCTCCTTGTAGCCCACCCATGAGAAGCAAGAGACCTTNAAGGCTTCCT
ATCCCACCATTACAG

AW339874

TTTTTTTTTTTTTCTGAGTAAGAACAGGCTTTATTTGTAAAACCACTCGTGACTCTTTAC
AAAGCAGGATACACAGAAGGGAAAAAAATACACAGGGCAAAATGGATGTTCTGAGTGCCA
CAAGGATCTGCTGAAAAAAGCCAAAGATGTAAGATGGCTGGGTATATATGAGAATGAATA
TTTCACTATATTCTGATTCAATTACCAGTCTCAGTGGCCCAGGATGAGCTTTTGGTGTGG
TCACATGGCCAACATTTGGATAACAAATGAGGAATAATGGTACCGCCTCACTAGTGCCTG
AGAACAGCATGTTCTGGAAAATGTCTCTGGAGTTAGAGATGTGTTAGCTTTTTCATTACA
GATGGAGAAATACAATGTTTACACAAC

BG399724

CATGATGTTCAGTATGATCAGTTAACCTTAACCTCTGAGCATCCTGAAGCAAAATCTAAA
TAATGCAGCTATTACCACTGGTGGTCCAGGCTCTGGTGAAGCCCTCTGAGCCCAGGAGGA
AGAGAAAGCATTGTCCAGAGGTAGGAACACAGTCTGGGAGCCCAGAGCTCTGGGAGGAGT
GGGAAAATGCTGCTTCCTGCTGCTTGCTTCTAGGCACCTGCTTCCGCCATCTCACTTACC
ATGGCTAGAGATGGGGGTGAGACTGGGGAAGGACAAAAGCAGGGAACAGATAAGGGATGG
AAATCAGAAGGGAATATAGAAAGAACTCTGGATGTGGAGAAATGCCGGTACCTGAGCATT
TTGTATCAATGGGAGTACCCTCTGTAACTGCTCAGTAGGTTACAAATGAAGAGTCCACCA
GTATTAGAAACAATTTAAACTTGCCAGTACCAACTGGGATGTGTGCCTTCAATTTGAAAA
TTGTATGTTTTATTTTTTAAATTTGTTAACAGCATTAATTTATAGAGTATTGATGTCATT
TATGTTTCTGAGGTGTTTCAA

BF475787

TCTGAGTAAGAACAGGCTTTATTTGTAAAACCACTCGTGACTCTTTACAAAGCAGGATAC
ACAGAAGGGAAAAAAATACACAGTGCAAAATGGATGTTCTGAGTGCCACAAGGATCTGCT
GAAAAAAGCCAAAGATGTAAGATGGCTGGGTATATATGAGAATGAATATTTCACTATATT
CTGATTCAATTACCAGTCTCAGTGGCCCAGGATGAGCTTTTGGTGTGGTCACATGGCCAA

CATTTGGATAACAAATGAGGAATAATGGTACCGCCTCACTAGTGCCTGAGAACAGCATGT
TCTGGAAAATGTCTCTGGAGTTAGAGATGTGTTAGCTTTTTCATTACAGATGGAGAAATA
CAATGTTTACACAACAGTCCAGGGGTGGGGTCAAAAGTTGGAAGGTGTCATTAGACGCAG
CCAAATAAAGTGAAGACAACCCAGGTGACTGGCAGCCCTGACTTGTGCGTGGGCGA

BF437145

CTGACTGTCCCGTTTTTATTTTTACCATTGAGCCTTCTACCAGTACTGAAATGGGCAAAA
GATGGCTGATAACAAATTACACTTTACCTGTGATGGTTACTCTATGCTAGTATCCTGTTT
TTTAAAAAATAGTTCTTATGAGGTGTTAAGAAAAGCTTTCGCTTGGATTCATACACAGTT
GACCCT

H64601

AGGAAGTTAAGAACAGTCCTAAAATCTCTTTGGCTTCTTTGTCCTGATATGCACCGGCAT
TTTCACAGTAGGAACTAGGGTTTCTGTCCAGTTTTTTTGGTTCTTTAAGGAATTAATGTT
ATTCTGGGTACAACTGCTTACATACATAGCACATATAGATGACATTTTTACAGGCCGTCT
TGTTAGACTGACATACATGGAGGATAGTGCCACCCGCCTCACAAGAACATCAGGTAAGCT
CAGGCACAGAGTCCNAGGGNATCTGTAAGGGCTTCGCCCACGCACAAGTCAGGGCTGCCA
GTCACCNGGGTTGTCTTCACTTTATTTGGGCTGCGTCTAATGACACCTTNCCAACTTTTT
GACCCCACCCTGGGGCTTGTTGTGTAAACCATTGTTATTTCTCCCNTCTGTAATGGAAAA
AGGTTAACACNTTTTTAACTTCCGGNGACATTTTTC

AF212365

gcacgagcga tgtcgctcgt gctgctaagc ctggccgcgc tgtgcaggag cgccgtaccc
cgagagccga ccgttcaatg tggctctgaa actgggccat ctccagagtg gatgctacaa
catgatctaa tccccggaga cttgagggac ctccgagtag aacctgttac aactagtgtt
gcaacagggg actattcaat tttgatgaat gtaagctggg tactccgggc agatgccagc
atccgcttgt tgaaggccac caagatttgt gtgacgggca aaagcaactt ccagtcctac
agctgtgtga ggtgcaatta cacagaggcc ttccagactc agaccagacc ctctggtggt
aaatggacat tttcctacat cggcttccct gtagagctga acacagtcta tttcattggg
gcccataata ttcctaatgc aaatatgaat gaagatggcc cttccatgtc tgtgaatttc
acctcaccag gctgcctaga ccacataatg aaatataaaa aaaagtgtgt caaggccgga
agcctgtggg atccgaacat cactgcttgt aagaagaatg aggagacagt agaagtgaac
ttcacaacca ctcccctggg aaacagatac atggctctta tccaacacag cactatcatc
gggttttctc aggtgtttga gccacaccag aagaaacaaa cgcgagcttc agtggtgatt
ccagtgactg gggatagtga aggtgctacg gtgcagctga ctccatattt tcctacttgt
ggcagcgact gcatccgaca taaaggaaca gttgtgctct gcccacaaac aggcgtccct
ttccctctgg ataacaacaa aagcaagccg ggaggctggc tgcctctcct cctgctgtct
ctgctggtgg ccacatgggt gctggtggca gggatctatc taatgtggag gcacgaaagg
atcaagaaga cttccttttc taccaccaca ctactgcccc ccattaaggt tcttgtggtt
tacccatctg aaatatgttt ccatcacaca atttgttact tcactgaatt tcttcaaaac
cattgcagaa gtgaggtcat ccttgaaaag tggcagaaaa agaaaatagc agagatgggt

ccagtgcagt ggcttgccac tcaaaagaag gcagcagaca aagtcgtctt ccttctttcc
aatgacgtca acagtgtgtg cgatggtacc tgtggcaaga gcgagggcag tcccagtgag
aactctcaag actcttcccc ttgcctttaa ccttttctgc agtgatctaa gaagccagat
tcatctgcac aaatacgtgg tggtctactt tagagagatt gatacaaaag acgattacaa
tgctctcagt gtctgcccca agtaccacct catgaaggat gccactgctt tctgtgcaga
acttctccat gtcaagtagc aggtgtcagc aggaaaaaga tcacaagcct gccacgatgg
ctgctgctcc ttgtagccca cccatgagaa gcaagagacc ttaaaggctt cctatcccac
caattacagg gaaaaaacgt gtgatgatcc tgaagcttac tatgcagcct acaaacagcc
ttagtaatta aaacatttta taccaataaa attttcaaat attgctaact aatgtagcat
taactaacga ttggaaacta catttacaac ttcaaagctg ttttatacat agaaatcaat
tacagtttta attgaaaact ataaccattt tgataatgca acaataaagc atcttcagcc
aaaaaaaaaa aaaaaa

AF208110

cggcgatgtc gctcgtgctg ataagcctgg ccgcgctgtg caggagcgcc gtaccccgag
agccgaccgt tcaatgtggc tctgaaactg ggccatctcc agagtggatg ctacaacatg
atctaatccc cggagacttg agggacctcc gagtagaacc tgttacaact agtgttgcaa
caggggacta ttcaattttg atgaatgtaa gctgggtact ccgggcagat gccagcatcc
gcttgttgaa ggccaccaag atttgtgtga cgggcaaaag caacttccag tcctacagct
gtgtgaggtg caattacaca gaggccttcc agactcagac cagaccctct ggtggtaaat
ggacattttc ctatatcggc ttccctgtag agctgaacac agtctatttc attggggccc
ataatattcc taatgcaaat atgaatgaag atggcccttc catgtctgtg aatttcacct
caccaggctg cctagaccac ataatgaaat ataaaaaaaa gtgtgtcaag gccggaagcc
tgtgggatcc gaacatcact gcttgtaaga agaatgagga gacagtagaa gtgaacttca
caaccactcc cctgggaaac agatacatgg ctcttatcca acacagcact atcatcgggt
tttctcaggt gtttgagcca caccagaaga aacaaacgcg agcttcagtg gtgattccag
tgactgggga tagtgaaggt gctacggtgc agctgactcc atattttcct acttgtggca
gcgactgcat ccgacataaa ggaacagttg tgctctgccc acaaacaggc gtcctttcc
ctctggataa caacaaaagc aagccgggag gctggctgcc tctcctcctg ctgtctctgc
tggtggccac atgggtgctg gtggcaggga tctatctaat gtggaggcac gaaaggatca
agaagacttc cttttctacc accacactac tgccccccat taaggttctt gtggtttacc
catctgaaat atgtttccat cacacaattt gttacttcac tgaatttctt caaaaccatt
gcagaagtga ggtcatcctt gaaaagtggc agaaaaagaa aatagcagag atgggtccag
tgcagtggct tgccactcaa aagaaggcag cagacaaagt cgtcttcctt ctttccaatg
acgtcaacag tgtgtgcgat ggtacctgtg gcaagagcga gggcagtccc agtgagaact
ctcaagacct cttccccctt gcctttaacc ttttctgcag tgatctaaga agccagattc
atctgcacaa atacgtggtg gtctacttta gagagattga tacaaaagac gattacaatg
ctctcagtgt ctgccccaag taccacttca tgaaggatgc cactgctttc tgtgcagaac
ttctccatgt caagcagcag gtgtcagcag gaaaaagatc acaagcctgc cacgatggct
gctgctcctt gtagcccacc catgagaagc aagagacctt aaaggcttcc tatcccacca
attacaggga aaaaacgtgt gatgatcctg aagcttacta tgcagcctac aaacagcctt
agtaattaaa acattttata ccaataaaat tttcaaatat tactaactaa tgtagcatta
actaacgatt ggaaactaca tttacaactt caaagctgtt ttatacatag aaatcaatta

cagctttaat tgaaaactgt aaccattttg ataatgcaac aataaagcat cttccaaaaa
aaaaaaaaaa aaaaaaaaaa aaaaaaaa

AF208111

cggcgatgtc gctcgtgctg ataagcctgg ccgcgctgtg caggagcgcc gtaccccgag
agccgaccgt tcaatgtggc tctgaaactg ggccatctcc agagtggatg ctacaacatg
atctaatccc cggagacttg agggacctcc gagtagaacc tgttacaact agtgttgcaa
caggggacta ttcaattttg atgaatgtaa gctgggtact ccgggcagat gccagcatcc
gcttgttgaa ggccaccaag atttgtgtga cgggcaaaag caacttccag tcctacagct
gtgtgaggtg caattacaca gaggccttcc agactcagac cagaccctct ggtggtaaat
ggacattttc ctatatcggc ttccctgtag agctgaacac agtctatttc attggggccc
ataatattcc taatgcaaat atgaatgaag atggcccttc catgtctgtg aatttcacct
caccaggctg cctagaccac ataatgaaat ataaaaaaaa gtgtgtcaag gccggaagcc
tgtgggatcc gaacatcact gcttgtaaga agaatgagga gacagtagaa gtgaacttca
caaccactcc cctgggaaac agatacatgg ctcttatcca acacagcact atcatcgggt
tttctcaggt gtttgagcca caccagaaga aacaaacgcg agcttcagtg gtgattccag
tgactgggga tagtgaaggt gctacggtgc aggtaaagtt cagtgagctg ctctggggag
ggaagggaca tagaagactg ttccatcatt cattgctttt aaggatgagt tctctcttgt
caaatgcact tctgccagca gacaccagtt aagtggcgtt catgggggtt ctttcgctgc
agcctccacc gtgctgaggt caggaggccg acgtggcagt tgtggtccct tttgcttgta
ttaatggctg ctgaccttcc aaagcacttt ttattttcat tttctgtcac agacactcag
ggatagcagt accattttac ttccgcaagc ctttaactgc aagatgaagc tgcaaagggt
ttgaaatggg aaggtttgag ttccaggcag cgtatgaact ctggagaggg gctgccagtc
ctctctgggc cgcagcggac ccagctggaa cacaggaagt tggagcagta ggtgctcctt
cacctctcag tatgtctctt tcaactctag tttttgaagt ggggacacag gaagtccagt
ggggacacag ccactcccca aagaataagg aacttccatg cttcattccc tggcataaaa
agtgntcaaa cacaccagag ggggcaggca ccagccaggg tatgatgggt actacccttt
tctggagaac catagacttc ccttactaca gggacttgca tgtcctaaag cactggctga
aggaagccaa gaggatcact gctgctcctt ttttgtagag gaaatgtttg tgtacgtggt
aagatatgac ctagcccttt taggtaagcg aactggtatg ttagtaacgt gtacaaagtt
taggttcaga ccccgggagt cttgggcatg tgggtctcgg gtcactggtt ttgactttag
ggctttgtta cagatgtgtg accaagggga aaatgtgcat gacaacacta gaggtagggg
cgaagccaga aagaagggaa gttttggctg aagtaggagt cttggtgaga ttttgctgtg
atgcatggtg tgaactttct gagcctcttg tttttcctca gctgactcca tattttccta
cttgtggcag cgactgcatc cgacataaag gaacagttgt gctctgccca caaacaggcg
tccctttccc tctggataac aacaaaagca agccgggagg ctggctgcct ctcctcctgc
tgtctctgct ggtggccaca tgggtgctgg tggcagggat ctatctaatg tggaggcacg
aaaggatcaa gaagacttcc ttttctacca ccacactact gccccccatt aaggttcttg
tggtttaccc atctgaaata tgtttccatc acacaatttg ttacttcact gaatttcttc
aaaaccattg cagaagtgag gtcatccttg aaaagtggca gaaaaagaaa atagcagaga
tgggtccagt gcagtggctt gccactcaaa agaaggcagc agacaaagtc gtcttccttc
tttccaatga cgtcaacagt gtgtgcgatg gtacctgtgg caagagcgag ggcagtccca
gtgagaactc tcaagacctc ttccccccttg cctttaacct tttctgcagt gatctaagaa

gccagattca tctgcacaaa tacgtggtgg tctactttag agagattgat acaaaagacg
attacaatgc tctcagtgtc tgccccaagt accacttcat gaaggatgcc actgctttct
gtgcagaact tctccatgtc aagcagcagg tgtcagcagg aaaaagatca caagcctgcc
acgatggctg ctgctccttg tagcccaccc atgagaagca agagacctta aaggcttcct
atcccaccaa ttacagggaa aaaacgtgtg atgatcctga agcttactat gcagcctaca
aacagcctta gtaattaaaa cattttatac caataaaatt ttcaaatatt actaactaat
gtagcattaa ctaacgattg gaaactacat ttacaacttc aaagctgttt tatacataga
aatcaattac agctttaatt gaaaactgta accattttga taatgcaaca ataaagcatc
ttccaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa

AF250309

atgtcgctcg tgctgctaag cctggccgcg ctgtgcagga gcgccgtacc ccgagagccg
accgttcaat gtggctctga aactgggcca tctccagagt ggatgctaca acatgatcta
atcccgggag acttgaggga cctccgagta gaacctgtta caactagtgt tgcaacaggg
gactattcaa ttttgatgaa tgtaagctgg gtactccggg cagatgccag catccgcttg
ttgaaggcca ccaagatttg tgtgacgggc aaaagcaact tccagtccta cagctgtgtg
aggtgcaatt acacagaggc cttccagact cagaccagac cctctggtgg taaatggaca
ttttcctata tcggcttccc tgtagagctg aacacagtct atttcattgg ggcccataat
attcctaatg caaatatgaa tgaagatggc ccttccatgt ctgtgaattt cacctcacca
ggctgcctag accacataat gaaatataaa aaaaagtgtg tcaaggccgg aagcctgtgg
gatccgaaca tcactgcttg taagaagaat gaggagacag tagaagtgaa cttcacaacc
actcccctgg gaaacagata catggctctt atccaacaca gcactatcat cgggttttct
caggtgtttg agccacacca gaagaaacaa acgcgagctt cagtggtgat tccagtgact
ggggatagtg aaggtgctac ggtgcagctg actccatatt ttcctacttg tggcagcgac
tgcatccgac ataaaggaac agttgtgctc tgcccacaaa caggcgtccc tttccctctg
gataacaaca aaagcaagcc gggaggctgg ctgcctctcc tcctgctgtc tctgctggtg
gccacatggg tgctggtggc agggatctat ctaatgtgga ggcacgaaag gatcaagaag
acttccttttt ctaccaccac actactgccc cccattaagg ttcttgtggt ttacccatct
gaaatatgtt tccatcacac aatttgttac ttcactgaat ttcttcaaaa ccattgcaga
agtgaggtca tccttgaaaa gtggcagaaa aagaaaatag cagagatggg tccagtgcag
tggcttgcca ctcaaaagaa ggcagcagac aaagtcgtct tccttctttc caatgacgtc
aacagtgtgt gcgatggtac ctgtggcaag agcgagggca gtcccagtga gaactctcaa
gacctcttcc cccttgcctt taacctttтc tgcagtgatc taagaagcca gattcatctg
cacaaatacg tggtggtcta ctttagagag attgatacaa aagacgatta caatgctctc
agtgtctgcc ccaagtacca cctcatgaag gatgccactg ctttctgtgc agaacttctc
catgtcaagc agcaggtgtc agcaggaaaa agatcacaag cctgccacga tggctgctgc
tccttgtagc ccacccatga gaagcaagag accttaaagg gttccttttc ccatcattta
caggggaaaa acgtgtgatg atc

AK095091

catattagag tctacagata tgcctttctt acagcaatcc tgcacccaca taaaagctac
attttcaata caagattaaa aggtattctg caaaatgtgc aaggttttca tgtctgctgg

tgtagctgta gtgatggctt catgaatttt tttctttttt gactatggtc cttacgctgg
attcatttat cttgaaatgg tgaacaatca cagctgcaga ccctcaattt atggtacata
tcaagcaatt tggctttttt tcttgtaatg aaaaaaaaaa gtttttttg ctttttttca
tgacactgct tcttgggagc actgccagca ttactagtgg cacttcgtat gggtcctaag
gtgttattga aggtttacga tattgcacta aacacgaaaa ataccagaga accactggag
atacttttta ctgtgatatg taatttactg gagacaggaa ctgctcgttt ggagatggtt
agcatcacag ggtgttttaa gtcgatactt gcaacccttg agctcaccac agtagcaaca
ggaggtggct aggaaattat tcacagcagg acagtacgca ctgcaattaa ttgtatgcag
ttatgattta ataccacatc tttatgctca cgtttctctc aactgtgaat ggtgccatgt
acagttggta tgtgtgtgtt taagttttga taaattttta acttttaata gttaaaatag
ttaactattg gtatggtagg aaatgataaa gtagactagt atctgtatac attttctgca
tttatgacat acctttttct tcatttttt caatatttta attgaaaagt tcatccgagt
ttcatctaag ttttttcaaa gtgatacaaa tctccaaaaa attttccaat atatgtattg
aaaaaatcca ggtgtaagtg gctctgcgca gtccaaacct gtgttgttca agggtcaact
gtgtatgaat ccaagcgaaa gcttttctta acacctcata agaactattt tttaaaaaac
aggaactagc atagagtaac catcacaggt aaagtgtaat ttgttatcag ccatcttttg
cccatttcag tactggtaga aggctcaatg gtaaaaataa aaacgggaca gtcagaagat
ctggaagtcc tgaccctgct ttcacctggc atgtgtaatc cagtcatgct cgtatcagtc
tctgtaggag cacttgaagg tattacataa atgctatcta actctgggaa acgccaacat
gtgattgcct ccagaggaat cttctttaaa aaaaaattca aaatgttatt tccttactag
gatgtcttta aagaattata acccttaccg tgcctccaca ttagatagat ccctgccacc
agcacccatg tggccaccag cagagacagc aggaggagag gcagccagcc tcccggcttg
cttttgtctg gaaaaaacaa agcttattca cctttggaaa acaaatccac acttatctct
taatttaaaa actaagactt ggtatacttt atagaggttt atttatttt tattattttt
tagttttgag acagagtctc gctttgttgc ctaggctgga gtgcagtggc gcaatctcgg
ttcactgcag cctccgtctc ccgggttcaa gcaatgctgc ctcagcctcc tgagtagctg
ggattacagg catgtgtcac cgcgcccagc cactttgtag agatttagat ccctttaaaa
ccatcagtca gaagctcttt agatagtctg ccaatcatat ctttttccct agagtgtgca
ggtcttgcat tagattctca aaagggatat gggacccagg aagttaagaa cagtcctaaa
atctctttgg cttctttgtc ctgatatgca ccggcatttt cacagtagga actagggttt
ctgtccagtt tttttggttc tttaaggaat taatgttatt ctgggtacaa ctgcttacat
acatagcaca tatagatgac atttttacag gccgtcttgt tagactgaca tacatggagg
atagtgccac ccgcctcaca agaacatcag gtaagctcag gcacagagtg cccaggaatc
tgtaaggctt cgcccacgca caagtcaggg ctgccagtca cctgggttgt cttcacttta
tttggctgcg tctaatgaca ccttccaact tttgacccca cccctggact gttgtgtaaa
cattgtattt ctccatctgt aatgaaaaag ctaacacatc tctaactcca gagacatttt
ccagaacatg ctgttctcag gcactagtga ggcggtacca ttattcctca tttgttatcc
aaatgttggc catgtgacca caccaaaagc tcatcctggg ccactgagac tagtaattga
atcagaatat agtgaaatat tcattctcat atatacccag ccatcttaca tctttggctt
ttttcagcag atccttgtgg cactcagaac atccattttg cactgtgtat ttttttccct
tctgtgtatc ctgctttgta aagagtcacg agtggtttta caaataaagc ctgttcttac
tcag

BM983744

TTTTTTTTTTTTTTTTCTGAGTAAGAACAGGCTTTATTTGTAAAACCACTCGTGACTCTT
TACAAAGCAGGATACACAGAAGGGAAAAAAATACACAGTGCAAAATGGATGTTCTGAGTG
CCACAAGGATCTGCTGAAAAAAGCCAAAGATGTAAGATGGCTGGGTATATATGAGAATGA
ATATTTCACTATATTCTGATTCAATTACCAGTCTCAGTGGCCCAGGATGAGCTTTTGGTG
TGGTCACATGGCCAACATTTGGATAACAAATGAGGAATAATGGTACCGCCTCACTAGTGC
CTGAGAACAGCATGTTCTGGAAAATGTCTCTGGAGTTAGAGATGTGTTAGCTTTTTCATT
ACAGATGGAGAAATACAATGTTTACACAACAGTCCAGGGGTGGGGTCAAAAGTTGGAAGG
TGTCATTAGACGCAGCCAAATAAAGTGAAGACAACCCAGGTGACTGGCAGCCCTGACTTG
TGCGTGGGCGAAGCCTTACAGATTCCTGGGCACTCTGTGCCTGAGCTTACCTGATGTTCT
TGTGAGGCGGGTGGCACTATCCTCCATGTATGTCAGTCTAACAAGACGGCCTGTAAAAAT
GTCATCTATATGTGCTATGTATGTAAGCAGTTGTACCCAGAATAACATTAATCCTCGTGC
CGAAT

CB305764

TTTTTTTTTTTTTTTGTTGGGCTGAAGATGCTTTATTATTGCATTATCAAAATGGTTATA
GTTTTCAATTAAAACTGTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGT
AGTTTCCAATCGTTAGTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTAT
AAAATGTTTTAATTACTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACA
CGTTTTTTCCCTGTAATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGT
GGGCTACAAGGAGCAGCAGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTG
CTGCTTGACATGGAGAAGTTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTT
GGGGCAGACACTGAGAGCATTGTAATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCAC
CACGTATTTGTGCAGATGAATCTGGCTTCTTAGATCACTGCAGAAAAGGTTAAAGGCAAG
GGGGAAGAGGTCTTGAGAGTTCTCACTGGGACTGCCCTCGCTCTTGCCACAGGTACCATC
GCACACACTGTTNACGTCATTGGAAAGAAGGAAGACGACTTTGTCTGCTGCCTTCTTTTG
AGTG

BM715988

TGGTTTTTGTTTTTTTTTCATTTTCTGTTGGATTACAGAAAAAGAATGGGACCCATTCAG
GTCTCGATTTCCAAAGGTAAAGATGGAAGGCTGGGCAGACTGGCTTTTGTTACCTGACAT
GCCGTAGGGTGAGCTTAGAGGAAGAAAGAAAACAATTTTTATTTGGCCAAAACAGAACAA
ATGCTGAAAAGGAAATCTTGTTTTTTTCCTAAAGCCAAATAGAAATGATTTGGGTATAAT
TTAAGAGTCCTTGTGTTGTACAGATATGGTGACTGATGTAGTTATTAATACTACCAACTT
AGTCATCAAGCCTCAATTTTCCTTTACCTGAAGGATTAAGTGAAAGCTTTTGGAGTTCAT
GATGTTCAGTATGATCAGTTAACCTTAACCTCTGAGCATCCTGAAGCAAAATCTAAATAA
TGCAGCTATTACCACTGGTGGTCCAGGCTCTGGTGAAGCCCTCTGAGCCCAGGAGGAAGA
GAAAGCATTGTCCAGAGGTAGGAACACAGTCTGGGAGCCCAGAGCTCTGGGAGGAGTGGG
AAAATGCTGCTTCCTGCTGCTTGCTTCTAGGCACCTGCTTCCGCCATCTCACTTACCATG
GCTAGAGATGGGGGTGAGACTGGGGAAGGACACAAGCAGGGAACAGATAAGGGATGGAAA
TCAGAAGGGAATATAGAAAGAACTCTGGATGTGGAGACATGCCGGTACCTGAGCATTTTG
TATCAATGGGAGTACCTCT

BM670929

TTTTTTTTTTTTTTTTTTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTACAG
TTTTCAATTAAAGCTGTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTA
GTTTCCAATCGTTAGTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATA
AAATGTTTTAATTACTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACAC
GTTTTTTTCCCTGTAATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGT
GGGCTACAAGGAGCAGCAGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTG
CTGCTTGACATGGAGAAGTTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTT
GGGGCAGACACTGAGAGCATTGTAATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCAC
CACGTATTTGTGCAGATGAATCTGGCTTCTTAGATCACTGCAGAAAAGGTTAAAGGCAAG
GGGGAAGAGGTCTTGAGAGTTCTCACTGGGACTTGCCTCGCTCTTGCCACAGGTACCATC
GCACACACTGTTGACGTCATTGGAAAGAAAGAAGACGACTTTGTCTGCTGCCTTCTT

BI792416

GCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTACAGTTTTCAATTAAAGCTGTAA
TTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTAGT

BI715216

CACGCGTCCGATTTTATACCAATAAAATTTTCAAATATTGCTAACTAATGTAGCATTAAC
TAACGATTGGAAACTACATTTACAACTTCAAAGCTGTTTTATACATAGAAATCAATTACA
GCTTTAATTGAAAACTGTAACCATTTTGATAATGCAACAATAAAGCATCTTCAGCCAAAA
AAAAAAA

N56060

AGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCATAAAAAAGAAGGCAGC
AGACAAAGTCGTCTTCCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGG
CAAGAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTTCCCCCCTTGCCTTTAACC
TTTTCTGCAGTGATCTAAGAAGCCAGATTCATCTGCACAAATACGTGGTGGTCTACTTTA
GAGAGATTGATACAAAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCA
TGAAGGATGCCACTGCTTTCTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTTTCAGCAG
G

CB241389

TTTTTTTTTTTTTTGTTTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTACAG
TTTTCAATTAAAGCTGTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTA
GTTTCCAATCGTTAGTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATA
AAATGTTTTAATTACTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACAC
GTTTTTTCCCTGTAATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTG
GGCTACAAGGAGCAGCAGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGC
TGCTTGACATGGAGAAGTTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTG

GGGCAGACACTGAGAGCATTGTAATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCACC
ACGTATTTGTGCAGATGAATCTGGCTTCTTAGATCACTGCAGAAAAGGTTAAAGGCAAGG
GGGAAGAGGTCTTGAGAGTTCTCACTGGGACTGCCCTCGCTCTTGCCACAGGTACCATCG
CACACACTGTTGACGTCATTGGAAAGAAGGAAGACGACTTTGTCTGCTGCCTTCTTTTGA
GTGGCAAGCCACTGCACTGGACCCATCTCTGCTATTTTCTTTTTCTNGCACTTTTCAAGG
ATGACTCACTTCTGCAATGGTTTTTGAGAATTCAGTGAAGTACAAATGTGTGATGGAACA
TAT

AV660618

CGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCCGACCG
TTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACTAGTGTTGCAACAGGGGACT
ATTCAATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCACACCAGAAGAAACAA
ACGCGAGCTTCAGTGGTGATTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTG
ACTCCATATTTTCCTACTTGTGGCAGCGACTGCATCCGACATAAAGGAACAGTTGTGCTC
TGCCCACAAACAGGCGTCCCTTTCCCTCTGGATAACAAC

BX088671

GCTGAGTGTGATGGTGTAAGCCTGTGGTCCCAGCTACTAGGGAGGCTGAGATGGGATTAC
AGGTGTGAGCCACGGCGCCTGGCCTAAAAGCATCTTTTTCTTTAACGCAGAGGTTATGTT
GTATTATTAGCATAAATGTTTTTTTCTGGGAATGCTTATTTCACACAGCACAATACTGAA
TCTTCTCTGGAATGTGGATCGATTTCAGATGGATGACTATTAAAATGTGTATATTTGCAG
ATTATCCTTAAAGGGCCACCTCATGCCTTCTAATTTATGTCTTACGGATAAAAAATCAAA
ATGAAGCATAAAGTAAAAACTGTGTCCAGCTTTACAAGTGGACGCTTAGTAATGGCTGAG
GCAATATGTTTAATGTAGCAAATTTTACTTATTTGTCATGATCAGTTTTCACAGTGCTTG
TAAGTGCTGGTAATAGAAGATGGACATGGTTTAGGTCAAAACTTGGACCAGAAACCAACT
TCCTTTGAAACAGCTCTACCAGNTATAAGAGCAATATG

CB154426

CTGTTGACGTCATTGGAAAGAAGGAAGACGACTTTGTCTGCTGCCTTCTTTTGAGTGGCA
AGCCACTGCACTGGACCCATCTCTGCTATTTTCTTTTTCTGCCACTTTTCAAGGATGACC
TCACTTCTGCAATGGTTTTGAAGAAATTCAGTGAAGTAACAAATTGTGTGATGGAAACAT
ATTTCAGATGGGTAAACCACAAGAACCTTAATGGGGGGCAGTAGTGTGGTGGTAGAAAAG
GAAGTCTTCTTGATCCTTTCTGTGAGAGGAGAAAAGCATTTGTTATCTGTGAACAGCAAA
CAGCAGGCTTTCACTCTGTAAACCATCCCTGACAAATGATCCCTTGCTAGAGAATGTCAG
CTGAGCACCAAGGGCCTTGTTAGTGACAGCAAGGAAAAACATCCTGATGTTCCTTTTGAA
CACATCACCTGAAACACACTGATGCTTAAACCTTAACTTTTTTTTTTTTGGAGACACAGT
CTCACTCTGT

CA434589

TTTTTTTTTTTTTTTTTTCTGAGTAAGAACAGGCTTTATTTGTAAAACCACTCGTGACTC
TTTACAAAGCAGGATACACAGAAGGGAAAAAAATACACAGTGCAAAATGGATGTTCTGAG
TGCCACAAGGATCTGCTGAAAAAAGCCAAAGATGTAAGATGGCTGGGTATATATGAGAAT
GAATATTTCACTATATTCTGATTCAATTACCAGTCTCAGTGGCCCAGGATGAGCTTTTGG
TGTGGTCACATGGCCAACATTTGGATAACAAATGAGGAATAATGGTACCGCCTCACTAGT
GCCTGAGAACAGCATGTTCTGGAAAATGTCTCTGGAGTTAGAGATGTGTTAGCTTTTTCA
TTACAGATGGAGAAATACAATGTTTACACAACAGTCCAGGGGTGGGGTCAAAAGTTGGAA
G

CA412162

TTTTTTTTTTTTTTTTTTGGCTGAAGATGCTTTATTGTTGCATTATCAAAATGGTTATAG
TTTTCAATTAAAACTGTAATTGATTTCTATGTATAAAACAGCTTTGAAGTTGTAAATGTA
GTTTCCAATCGTTAGTTAATGCTACATTAGTTAGCAATATTTGAAAATTTTATTGGTATA
AAATGTTTTAATTACTAAGGCTGTTTGTAGGCTGCATAGTAAGCTTCAGGATCATCACAC
GTTTTTTCCCTGTAATTGGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTG
GGCTACAAGGAGCAGCAGCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGC
TGCTTGACATGGAGAAGTTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACGTG
GGGCAGACACTGAGAGCATTGTAATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCACC
ACGTATTTGTGCAGATGAATCTGGCTTCTTAGATCACTGCAGAAAAGGTTAAAGGCAAGG
GGGAAGA

CA314073

TTTTTTTTTTTTTTTTTTGAAAGGGTCAGGACTTCCAGATCTTCTGACTGTCCCGTTTTT
ATTTTTACCATTGAGCCTTCTACCAGTACTGAAATGGGCAAAAGATGGCTGATAACAAAT
TACACTTTACCTGTGATGGTTACTCTATGCTAGTTCCTGTTTTTTAAAAAATAGTTCTTA
TGAGGTGTTAAGAAAAGCTTTCGCTTGGATTCATACACAGTTGACCCTTGAACAACACAG
GTTTGGACTGCGCAGAGCCACTTACACCTGGATTTTTTCAATACATATATTGGAAAATTT
TTTGGAGATTTGTATCACTTTGAAAAAACTTAGATGAAACTCGGATGAACTTTTCAATTA
AAATATTGAAAAAAATGAAGAAAAAGGTATGTCATAAATGCAGAAAATGTATACAGATAC
TAGTCTACTTTATCATTTCCTACCATACCAATAGTTAACTATTTTAACTATTAAAAGTTA
AAAATTTATCAAAACTTAAACACACACATACCAACTGTACATGGCACCATTCACAGTTGA
GAGAAACGTGAGCATAAAGATGTGGTATTAAATCATAACTGCATACAATTAATTGCAGTG
CGTACTGTCCTGCTGTGAATATTTCCTAGCCCTCGTGCCGAATC

BF921554

GTGGGTGACCGTGGCTTGCCACTCAAAAGAAGGCAGCAGACAAAGTCGTCTTCCTTCTTT
CCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCCAGTG
AGAACTCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAAGCC
AGATTCATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACGATT
ACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTTCTGTG
CATAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG

ATGGCTGCTGCTCCTTGTAGCCCACCCATGAGAAGCAAGAGACCTTAAAGGCTTCCTATC
CCACCAATTACAGGGAAAAAAACGTGTGATGATCCTGAAGCCACGGTCAA

BF920093

TAGAGGATCCCGGTCGACGGTGGTTCAGTGATCATCACACTTTTTCCCTGTAATAGGTGG
GATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCAGCCAT
CGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAAGTTATG
CACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGCATTGT
AATCGTCTTTTGTATCAATCTCTCTAAAGTAGACCACCACGTATTTGTGCAGATGAATCT
GGCTTCTTAGATCACTGCAGAAAAGGTTAAAGGCAAGGGGGAAGAGGTCTTGAGAGTTCT
CACTGGGACTGCCCTCGCTCTTGCCACAGGTACCATCGCACACACTGTTGACGTCATTGG
AAAGAAGGAAGACGACTTTGTCTGCTGCCTTCTTTTGAGTGGCAAGCCACGGTCAACCCA
CAAGCCACGGTCAACCCAC

AV685699

TCTACGTGGTAAGATATGACCTAGCCCTTTTAGGTAAGCGAACTGGTATGTTAGTAACGT
GTACAAAGTTTAGGTTCAGACCCCGGGAGTCTTGGGCATGTGGGTCTCGGGTCACTGGTT
TTGACTTTAGGGCTTTGTTACAGATGTGTGACCAAGGGGAAAATGTGCATGACAACACTA
GAGGTAGGGGCGAAGCCAGAAAGAAGGGAAGTTTTGGCTGAAGTAGGAGTCTTGCGACTG
CATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCTGGA
TAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTCTGCTGGTGGC
CACATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGAC
TTCCTTTTCTACCACCACACTACTGCCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGA
AATATGTTTCCATCACACAATTTGTTACTTCACTGAATTTCTTCAAAACCATTGCAGAAG
TGAGGTCATCCTTGAAAGTGGCAGAGTAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCAC
TCGTGCGATGGTCTT

AV650175

GGCACGAGCACTGGCTGAAGGAAGCCAAGAGGATCACTGCTGCTCCTTTNTTCTAGAGGA
AATGTTTGTCTACGTGGTAAGATATGACCTAGCCCTTTTAGGTAAGCGAACTGGTATGTT
AGTAACGTGTACAAAGTTTAGGTTCAGACCCCGGGAGTCTTGGGCATGTGGGTCTCGGGT
CACTGGTTTTGACTTTAGGGCTNTGTTACAGATGTGTGACCAAGGGGAAAATGTGCATGA
CAACACTAGAGCTGACTCCATATTTTCCTACTTGTGGCAGCGACTGCATCCGACATAAAG
GAACAGTTGTGCTCTGCCCACANACAGGCGTCCCTTTCCCTCTGGATAACAACATAAGCA
AGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTCTGCTGGTGGCACATGGGTGCTGGT
GGAGGGATCTATCTAATGTGGAGGCACGGATCAAGAAGACTTNCTTNTCTACCACCACAC
TACTGGCCCCAATAAGGGTCTNGTGGNTACCCCATCTGAATATGTTCATACACAATTTGT
ACTCACTGAATTCTCAAAACATTGAGAGTGAGGCATCCTGAAAGTGCGAAAAGANATGCN
AATGGTCAGTGCATGCTGCACTAGCAGCATGGACTT

BX483104

GATCCCGCGCAGTGGCCCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAG
GAGCGCCGTACCCCGAGAGCCGACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGA
GTGGATGCTACAACATGATCTAATCCCCGGAGACTTGAGGGACCTCCGAGTAGAACCTGT
TACAACTAGTGTTGCAACAGGGGACTATTCAATTTTGATGAATGTAAGCTGGGTACTCCG
GGCAGATGCCAGCATCCGCTTGTTGAAGGCCACCAAGATTTGTGTGACGGGCAAAAGCAA
CTTCCAGTCCTACAGCTGTGTGAGGTGCAATTACACAGAGGCCTTCCAGACTCAGACCAG
ACCCTCTGGTGGTAAATGGACATTTTCCTACATCGGCTTCCCTGTAGAGCTGAACACAGT
CTATTTCATTGGGGCCCATAATATTCCTAATGCAAATATGAATGAAGATGGCCCTTCCAT
GTCTGTGAATTTCACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAAAAGTG
TGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTAAGAAGAATGAGGAGAC
AGTAGAAGTGAACTTCACAACCACTCCCCTGGGAAACAGATACATGGCTCTTATCCAACA
CAGCACTATCATTCGG

CD675121

GTCTTGCATTAGATTCTCAAAAGGGATATGGGACCCAGGAAGTTAAGAACAGTCCTAAAA
TCTCTTTGGCTTCTTTGTCCTGATATGCACCGGCATTTTCACAGTAGGAACTAGGGTTTC
TGTCCAGTTTTTTTGGTTCTTTAAGGAATTAATGTTATTCTGGGTACAACTGCTTACATA
CATAGCACATATAGATGACATTTTTACAGGCCGTCTTGTTAGACTGACATACATGGAGGA
TAGTGCCACCCGCCTCACAAGAACATCAGGTAAGCTCAGGCACAGAGTGCCCAGGAATCT
GTAAGGCTTCGCCCACGCACAAGTCAGGGCTGCCAGTCACCTGGGTTGTCTTCACTTTAT
TTGGCTGCGTCTAATGACACCTTCCAACTTTTGACCCCACCCCTGGACTGTTGTGTAAAC
ATTGTATTTCTCCATCTGTAATGAAAAAGCTAACACATCTCTAACTCCAGAGACATTTTC
CAGAACATGCTGTTCTCAGGCACTAGTGAGGCGGTACCATTATTCCTCATTTGTTATCCA
AATGTTGGCCATGTGACCACACCAAAAGCTCATCCTGGGCCACTGAGACTGGTAATTGAA
TCAGAATATAGTGAAATATTCATTCTCATATATACCCAGCCATCTTACATCTTTGGCTTT
TTTCAGCAGATCCTTGTGGCACTCAGAACATCCATTTTGCACTGTGTATTTTTTTCCCTT
CT

BE081436

TGTGTAACTCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAA
GCCAGATTCATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACG
ATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGGAGGATGCCACTGCTTTCT
GTGCAGAACTTCTCCATGTCAAGTAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCC
ACGATGGCTGCTGCTCCTTGTAGCCCACCCATGAGAAGCAAGAGACCTTAAAGGCTTCCT
ATCCCACCAATTACAGGGAAAAAACGTGTGATGAT

AW970151

CTGAAATATGTTTCCATCACACAATTTGTTACTTCACTGAATTTCTTCAAAACCATTGCA
GAAGTGAGGTCATCCTTGAAAAGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGC
AGTGGCTTGCCACTCAAAAGAAGGCAGCAGACAAAGTCGTCTTCCTTCTTTCCAATGACG
TCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCCAGTGAGAACTCTC

AAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAAGCCAGATTCATC
TGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACGATTACAATGCTC
TCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTTCTGTGCAGAACTTC
TCCATGTCAAGTAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACGATGGCTGCT
GCTCCTTGTAGCCCACCCATGAGAAGCAAGAGACCTTAAAGGCTTCCTATCCCACCAATT
ACAGGGAAAAAAACGTGTGATGATCCCTGAAGCTTACTATGCAGCCTACANACAGCCTTA
GTAATAAAACATTTTATCCAATAAAATTTCAAATTTTGCTTAACTATGTGCATAAACTAC
GATTGAAAACTCTTTACACT

AW837146

CATTGTGGTTGCAGCTGCATAGTAAGCTTCAGGATCATCACACGTTTTTTCCCTGTAATT
GGTGGGATAGGAAGCCTTTAAGGTCTCTTGCTTCTCATGGGTGGGCTACAAGGAGCAGCA
GCCATCGTGGCAGGCTTGTGATCTTTTTCCTGCTGACACCTGCTGCTTGACATGGAGAAG
TTCTGCACAGAAAGCAGTGGCATCCTTCATGAGGTGGTACTTGGGGCAGACACTGAGAGC
ATTGTAATCGTCTTTTGTATCAATCTCCCTAAAGTAGACCACCACGTATTTGTGCAGATG
AATCTGGCTTCTTAGATCACTGCAGAAAAGGTTAAAGGCAAGGGGGAAGAGGTCTTGAGA
GTTCTCACTGGGACTGCCCTCGCTCTTGCCACAGGTACCATCGCACACACTGTTGACGTC
ATTGGAAAGAAGGAAGACGACTTTGTCTGCTGCCTTCTTTTGAGTGGCAAGCCACTGCAC
TGGACCCATCT

AW368264

GTGAATAAGCTTTGTTTTTTCCAGACAAAAGCAAGCCAGGAGGCTGGCTGCCTCTCCTCC
TGCTGTCTCTGCTGGTGGCCACATGGTTGCTGGTGGCAGGGATCTATCTAATGTGGAGGC
ACGGTAAGGGTTATAATTCTTTAAAGTCATCCTAGTAAGGAAATAACATTTGGAATTTTT
TTTTAAAGAAGATTCCTCTGGAGGCAATCACCTGTTGGCGTTTCCCAGAGTTAGATAGCA
TTTATGTAATACCTTCAAGTGCTCCTACAGAGACTGATACGAGCATGACTGGATTACACA
TGCCAGGTGAAAGCAGGGCCAGGACTTCCAGATCTTCTGACTGTCCCGTTTTTATTTTTA
CCATTGAGCCTTCTACCAGAACTGAAATGGGCAAAAGATGGCTGATAACAAATTACACTT
TACCTGTGATGGTTACTCTATGCTAGTTCCTGTTTTTAAAAAAATAGTTCTTATGAGGTG
TCAAGAAAAGCTTTCGCTTGGATTCATACACAGTTGACCCTTGAACAACACAG

D25960

GATCCTGAAGCTTACTATGCAGCCTACAAACAGCCTTAGTAATTAAAACATTTTATACCA
ATAAAATTTTCAAATATTGCTAACTAATGTAGCATTAACTAACGATTGGAAACTACATNN
ACAACTTCAAAGCTGTTTTATACATAGAAATCAATTACAGCTTTAATTGAAAACTATAAC
CATTTTGATAATGCAACANTAAAGCATCTTCAGCCAAA

AV709899

GCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAATTACACAGAGGCCTTCCAGACTCAGA
CCAGACCCTCTGGTGGTAAATGGACATTTTCCTATATCGGCTTCCCTGTAGAGCTGAACA

CAGTCTATTTCATTGGGGCCCATAATATTCCTAATGCAAATATGAATGAAGATGGCCCTT
CCATGTCTGTGAATTTCACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTAAGAAGAATGAGG
AGACAGTAGAAGTGAACTTCACAACCACTCCCCTGGGAAACAGATACATGGCTCTTATCC
AACACAGCACTATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGC
GAGCTTCAGTGGTGATTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTC
CATATTTTCCTACTTGTGGCAGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCC
CACAAACAGGCGTNCCTTTTCCTCTGGATAACAACAAAAGCAAGCCGGGAGGCTTGGCTG
CTCTCCTTCTGCTGGCCTTTGCTGTGGCCACATTGGTGCTGGTGGCAGGGATCTATCTAA
TGTGGATGCACGTCTCGTGGTTTACCCATCTGAAATATGTTCN

BX431018

ATTTTTCCTCTTGTGGCAGCGACTGGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCA
CAAACAGGCGTCCCTTTCCCTCTGGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCT
CTCCTCCTGCTGTCTCTGCTGGTGGCCACATGGGTGCTGGTGGCAGGGATCTATCTAATG
TGGAGGCACGAAAGGATCAAGAAGACTTCCTTTTCTACCACCACACTACTGCCCCCCATT
AAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCATCACACAATTTGTTACTTCACT
GAATTTCTTCAAAACCATTGCAGAAGTGAGGTCATCCTTGAAAAGTGGCAGAAAAAGAAA
ATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAAGAAGGCAGCAGACAAAGTC
GTCTTCCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAG
GGCAGTCCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGT
GATCTAAGAAGCCAGATTCATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGAT
ACAAAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCC
ACTGCTTTCTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCA
CAAGCCTGCCACGATGGCTGCTGCTCCTTGTAGCCCACCCATGAGAAGCAAGAGACCTTA
AGGCTTCTATCCCACCANTACAGGNAAAAACGTGTGATGATCCTGAAGCTTACTATGCAG
CCTACAACAGGCTTAGTATTAAAACATTTATACCCATAAATTTTCAAATTGCT

AL535617

TAGGTGACACTATAGAACAAGTTTGTACAAAAAAGCAGGCTGGTACCGGTCCGGAATTCC
CGGGATAGTGGMCCGGCGAKGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGC
GCCGTACCCCGAGAGCCGACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCARAGTGG
ATGSKACAACATGATCTAATCCCGGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACA
ACTAGTGTTGCAACAGGGGACTATTCAATTTTGATGAATGTAAGCTGGGTACTCCGGGSA
GATGCCAGCATCCGCTTGTTGAAGGCCACCAAGATTTGTGTGAMGGGCAAAAGCAACWTC
CAGTCCTACAGCWGTGTGAGGTAGCAATTACACAGAGAGCACATATCCAGACTCTAGACC
AGACCCTCTGGWGGTAAATGGACATTTTCCTATATCGGCTTCCCTGTAGAGCTGAACACA
GTCTATATTCATTGGGGCCCAWAATAWWCCTAATGCAAATATGAATGAAGATGGCCCTTC
CATGTCTGTGAATTTCACCTCACCAGGCTGCCTAGACCACATAATGAAATAWAAAAAAAA
GTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTAAGAAGAATGARGA
GACAGTAGAAGTGAACTTCACAACCACTCCCCTGGGAAACAGATAMATKGCTCTTATCCA
ACACARMACTATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCG

AGCTTCAGTGGTGATTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCC
ATATTTTCCTACTTGTGGCAGCGWCTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCC
ACAAACAGGCGTCCCTTTYCCTCTGGATAACAACAAAAGCAACYGGGAGSTGGYTGYCT

AL525465

WAATWAKADDRATANHTGAAAACTATAACCATTTNTGATAATNGNAANAATAAAGCATCT
TCAGCCAAACATCTAGTCTTCCATAGACCATGCATTGCAGTGTACCCAGAWCTGTTTAGC
TAATATTCTATGTTTAATTAATGAATACTAACTCTAAGAACCCCTCACTGATTCACTCAA
TAGCATCTTAAGTGAAAAACCTTCTATTACATGCAAAAAATCATTGTTTTTAAGATAACA
AAAGTAGGGAATAAACAAGCTGAACCCACTTTTACTGGACCAAATGATCTATTATATGTG
TAACCACTTGTATGATTTGGTATTTGCATAAGACCTTCCCTCTACAAACTAGATTCATAT
CTTGATTCTTGTACAGGTGCCTTTTAACATGAACAACAAAATACCCACAAACTTGTCTAC
TTTTGCCTAAAGTTACCTATTAGAGGTCACTGTSAGAGTKCTCAGTTTCTTAGTTACTAT
TTAASTTTTSATGTTCAAAATGAAAATAATTCTKAAGTKGAAAGSGCTCTTGAAGTAACC
TTTTTATAAATGAGTTATTATAATGGTTTACTTAAATAAAAVAGAGGGGKTTTTGCGGTG
GCTCATGCCTCCAATCCCAGCACTTTGGCAAGGCCAAGGCAAAAVGATCGCTCAAGACCA
GGCTACGTCACAAAGCGAGACCTCCATCTCTACAAAAGATTTAAAAAATTAGCTGAGTGT
GATGGTGTGAGCCTGTGGTCCCAGCTACTAGGGAGGCTGAGATGGGAGGATCACTTGAGC
CCTGGAGGTCAAGGGTGCAGTAAACGGTGATTGTGCCACTGCACTCCATCCTGGGTGAGA
GCAGACCCTGTCTAAAACAAACAAACGAAAAAACCCCCACAGAATGACAGAACATAAAAG
ATGCACATTTTGTCTTCCAACTTTTTACTCTTCTAAAAGCATCTTTTTTAAATTTTTTAA
ATTTTTTTTTTTTTGAGACAGAGTTTCACTCTGTCACACAGGCTGGAGTGMGTGGCGTGA
CTCGGCTCACTAMAACTCTGCYTCCGGGGTYACSCATCTCCTGCWCAGCTCCTGAGAAGC
KGGAYAMAGGMCCACACAAACCAGTAAYTTTATWTTTTGAAAAAGGGTTYACCTGTASMA
GRAGGCTGAATCCGACMAARTMACCMCCACYYCAAADGAGGAWAAGKGKRSMGGSCBGGC
A

BX453536

TTATGGGGGGCAGTAGTGTGGTGGTAGAAAAGGAAGTCTTCTTGATCCTTTCGTGCCTCC
CATTAGATAGATCCCTGCCACCAGCACCCATGTGGCCACCAGCAGAGACAGCAGGAGGAG
AGGCAGCCAGCCTCCCGGCTTGCTTTTGTTGTTATCCAGAGGGAAAGGGACGCCTGTTTG
TGGGCAGAGCACAACTGTTCCTTTATGTCGGATGCAGTCGCTGCCACAAGTAGGAAAATA
TGGAGTCAGCTGCACCGTAGCACCTTCACTATCCCCAGTCACTGGAATCACCACTGAAGC
TCGCGTTTGTTTCTTCTGGTGTGGCTCAAACACCTGAGAAAACCCGATGATAGTGCTGTG
TTGGATAAGAGCCATGTATCTGTTTCCCAGGGGAGTGGTTGTGAAGTTCACTTCTACTGT
CTCCTCATTCTTCTTACAAGCAGTGATGTTCGGATCCCACAGGCTTCCGGCCTTGACACA
CTNTNTTTTATATTTCATTATGTGGTCTAGGCAGCCTGGTGAGGTGAAATTCACAGACAT
GGAAGGGCCATCTTCATTCATATTTGCATTAGGAATATTATGGGCCCCAATGAAATAGAC
TGTGTTCAGCTCTACAGGGGAAGCCGATATAGGAAAATGTCCATTTACCACCAGAGGGTC
TGGTCTGAGTCTTGAAGGCCTTTTGTGTTATTGCACCTTACACAGCTGTTAGACTGGGAA
GTTGCTTTTGCCCCGCACACAAATCTTGTGGGCCTTCAACAGCGGATGCTGCCATTTGCC
CCGAAGTCCCCAGCTCAATTCATTAAAAATTGAATAGGCCCCTTGTGGCAACCCTAGTTG

GTACAGGGTTTTACTTGGGGGGCCCCTCTAAGTTTCCCCGGGATATAAACAAAGTGTGG

BX453537

TTATGGGGGGCAGTAGTGTGGTGGTAGAAAAGGAAGTCTTCTTGATCCTTTCGTGCCTCC
ACATTAGATAGATCCCTGCCACCAGCACCCATGTGGCCACCAGCAGAGACAGCAGGAGGA
GAGGCAGCCAGCCTCCCGGCTTGCTTTTGTTGTTATCCAGAGGGAAAGGGACGCCTGTTT
GTGGGCAGAGCACAACTGTTCCTTTATGTCGGATGCAGTCGCTGCCACAAGTAGGAAAAT
ATGGAGTCAGCTGCACCGTAGCACCTTCACTATCCCCAGTCACTGGAATCACCACTGAAG
CTCGCGTTTGTTTCTTCTGGTGTGGCTCAAACACCTGAGAAAACCCGATGATAGTGCTGT
GTTGGATAAGAGCCATGTATCTGTTTCCCAGGGGAGTGGTTGTGAAGTTCACTTCTACTG
TCTCCTCATTCTTCTTACAAGCAGTGATGTTCGGATCCCACAGGCTTCCGGCCTTGACAC
ACTTTTTTTTATATTTCATTATGTGGTCTAGGCAGCCTGGTGAGGTGAAATTCACAGACA
TGGAAGGGCCATCTTCATTCATATTTGCATTAGGAATATTATGGGCCCCAATGAAATAGA
CTGTGTTCAGCTCTACAGGGAAGCCGATATAGGAAAATGTCCATTTACCACCAGAGGGTC
TGGTCTGAGTCTGGAAGGCCTCTGTGTAATTGCACCTCACACAGCTGTAGGACTGGGAGT
TGCTTTTGCCCGTACACAAATCTTGTTGGCCTTCAACAAGCGGATGCTGGCATCTGGCGG
GGGTACCCAGCTTACATTCATCAAAATTGAATAGTCCCCTTGTTGCAACACTAGTTTGTA
AACAGGTTCTACTCCGGGGGTCCCCTCAGTCTCCCGG

AV728945

CAAATATGAATGAAGATGGCCCTTCCATGTCTGTGAATTTCACCTCACCAGGCTGCCTAG
ACCACATAATGAAATATAAAAAAAAGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACA
TCACTGCTTGTAAGAAGAATGAGGAGACAGTAGAAGTGAACTTCACAACCACTCCCCTGG
GAAACAGATACATGGCTCTTATCCAACACAGCACTATCATCGGGTTTTCTCAGGTGTTTG
AGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGATTCCAGTGACTGGGGATAGTG
AAGGTGCTACGGTGCAACTGACTCCATATTTTCCTACTTGTGGCAGCGACTGCATCCGAC
ATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCTGGATAACAAC

AV728939

GCAAATATGAATGAAGATGGCCCTTCCATGTCTGTGAATTTCACCTCACCAGGCTGCCTA
GACCACATAATGAAATATAAAAAAAAGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAAC
ATCACTGCTTGTAAGAAGAATGAGGAGACAGTAGAAGTGAACTTCACAACCACTCCCCTG
GGAAACAGATACATGGCTCTTATCCAACACAGCACTATCATCGGGTTTTCTCAGGTGTTT
GAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGATTCCAGTGACTGGGGATAGT
GAAGGTGCTACGGTGCAGCTGACTCCATATTTTCCTACTTGTGGCAGCGACTGCATCCGA
CATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCTGGATAACAAC

AV727345

GCAAATATGAATGAAGATGGCCCTTCCATGTCTGTGAATTTCACCTCACCAGGCTGCCTA
GACCACATAATGAAATATAAAAAAAAGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAAC

ATCACTGCTTGTAAGAAGAATGAGGAGACAGTAGAAGTGAACTTCACAACCACTCCCCTG
GGAAACAGATACATGGCTCTTATCCAACACAGCACTATCATCGGGTTTTCTCAGGTGTTT
GAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGATTCCAGTGACTGGGGATAGT
GAAGGTGCTACGGTGCAGCTGACTCCATATTTTCCTACTTGTGGCAGCGACTGCATCCGA
CATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCTGGATAACAAC
AAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTCTGCTGGTGGCCACATGG
GTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
TTTACCACCACACTACTGTCTCCCATTAAAGATCTTGTGGTTTATCCATCTGAAATATTG
TTCCATTACACATATTGGTACCTAACTGAAATTCTTTAAAACCATTGCAAATTGAGGTCA
CTCTTGAAAGGGCGTG

Sequences identified as those of CACNA1D cluster

BM128550

CGGCTCCTACCTTTTGCCCGATCCCCTTCCCCATTCCGCCCCCGCCCCAACGCAGTGCAC
AGTGCCCTGCACACAGTAGTCGCTCAATAAATGTTCGTGGATGATGATGATGATGATGAT
GAAAAAAATGCAGCATCAACGGCAGCAGCAAGCGGACCACGCGAACGAGGCAAACTATGC
AAGAGGCACCAGACTTCCTCTTTCTGGTGAAGGACCAACTTCTCAGCTGAATAGCTCCAA
GCAAACTGTCCTGTCTTGGCAAGCTGCAATCGATGCTGCTAGACAGGCCAAGGCTGCCCA
AACTATGAGCACCTCTGCACCCCCACCTGTAGGATCTCTCTCCCAAAGAAAACGTCAGCA
ATACGCCAAGAGCAAAAAACAGGGTAACTCGTCCAACAGCCGACCTGCCCGCGCCCTTTT
CTGTTTATCACTCAATAACCCCATCCGAAGAGCCTGCATTAGTATAGTGGAATGGAAACA
TTTGACATATTTATATTATTGGCTATTTTTTGCCAAT

BI755471

GAATATGACCCTGAGGCAAAGGGAAGGATAAACACCTTGATGTGGTCACTCTGCTTCGAC
GCATCCAGCCTCCCCTGGGGTTTGGGAAGTTATGTCCACACAGGGTAGCGTGCAAGAGAT
TAGTTGCCATGAACATGCCTCTCAACAGTGACGGGACAGTCATGTTTAATGCAACCCTGT
TTGCTTTGGTTCGAACGGCTCTTAAGATCAAGACCGAAGGGAACCTGGAGCAAGCTAATG
AAGAACTTCGGGCTGTGATAAAGAAAATTTGGAAGAAAACCAGCATGAAATTACTTGACC
AAGTTGTCCCTCCAGCTGGTGATGATGAGGTAACCGTGGGGAAGTTCTATGCCACTTTCC
TGATACAGGACTACTTTAGGAAATTCAAGAAACGGAAAGAACAAGGACTGGTGGGAAAGT
ACCCTGCGAAGAACACCACAATTGCCCTACAGGCGGGATTAAGGACACTGCATGACATTG
GGCCAGAAATCCGGCGTGCTATATCGTGTGATTTGCAAGATGACGAGCCTGAGGAAACAA
AACGAGAAGAAGAAGATGATGTGTTCAAAAGAAATGGTGCCCTGCTTGGAAACCATGTCA
ATCATGTTAATAGTGATAGGAGAGATTCCCTTCAGCAGACCAATAGCACCACCGTCCCCT
GCATTGTCCAAAGGCCTTCAATTCCACCTGCAAGTGATACTGAGAAACCGCTGTTTCCTC
CAGCAGGAAATTCGGGGTGTCATAACCATCATAACCATTAATTCCATAGGAAAGCAAGGT
TCCCACTTCAACAATGCCAGTCTCGAATAGTGCCAATATGTCCAAAGCTTGCCATGGTAA
GCGGGCCAGCATTGGGAACC

BQ549084

GCACGAGATTAATTAGACTTTTGTATAAGAGATGTCATGCCTCAAGAAAGCCATAAACCT
GGTAGGAACAGGTCCCAAGCGGTTGAGCCTGGCAGAGTACCATGCGCTCGGCCCCAGCTG
CAGGAAACAGCAGGCCCCGCCCTCTCACAGAGGATGGGTGAGGAGGCCAGACCTGCCCTG
CCCCATTGTCCAGATGGGCACTGCTGTGGAGTCTGCTTCTCCCATGTACCAGGGCACCAG
GCCCACCCAACTGAAGGCATGGCGGCGGGGTGCAGGGGAAAGTTAAAGGTGATGACGATC
ATCACACCTGTGTCGTTACCTCAGCCATCGGTCTAGCATATCAGTCACTGGGCCCAACAT
ATCCATTTTTAAACCCTTTCCCACAAATACACTGCGTCCTGGTTCCTGTTTAGCTGTTCT
GAAATACGGTGTGTAAGTAAGTCAGAACCCAGCTACCAGTGATTATTGCGAGGGCAATGG
GACCTCATAAATAAG

BQ549571

TTTTTTTTTTTTTTTTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCA
GTTCAAATACAATACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATA
GTATATTACAAGTCATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTT
TACCTGGTTGCGAGTGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGA
AACCGACCATCGGAGTGATATTCTCTTATGTAAACTGGCGTCACATCACAGAAAACCTTA
TTTATGAGGTCCCATTGCCCTCGCAATAATCACTGGTAGCTGGGTTCTGACTTACTTACA
CACCGTATTTCAGAACAGCTAAACAGGAACCAGGACGCAGTGTATTTGTGGGAAAGGGTT
TAAAAATGGATATGTTGGGCCCAGTGACTGATATGCTAGACCGATGGCTGAGGTAACGAC
ACAGGTGTGATGATCGTCATCACCTTTAACTTTCCCCTGCACCCCGCCGCCATGCCTTCC
AGTTGGGTGGGCCTGGT

AI693324

CTCTGAGCACTACAATCAGCCAGATTGGTTGACACAGATTCAAGATATTGCCAACAAAGT
CCTCTTGGCTCTGTTCACCTGCGAGATGCTGGTAAAAATGTACAGCTTGGGCCTCCAAGC
ATACTCTTGTTCTCTTTACAACCGGTTTGATTGCTTCGTGGTGTGTGGTGGAATCACTGA
GACGATCTTGGTGGAACTGGAAATCATGTCTCCCCTGGGGATCTCTGTGTTTCGGTGTGT
GCGCCTCTTAAGAATCTTCAAAGTGACCAGGCACTGGACTTCCCTGAGCAACTTAGTGGC
ATCCTTATTAAACTCCATGAAGTCCATCGCTTCGCTGTTGCTTCTGCTTTTTCTCTTCAT
TATCATCTTTTCCTTGCTTGGGATGCAGCTGTTTGGCGGCAAGTTTAATTTTGATG

R25307

ACCAGCAGACCTGACTGTCCCCAGCAGCTTCCGGAACAAAAACAGCGACAAGAGAGGAGT
GCGGACAGTTGGTGGAGGCAGTCCTGATATCCGAAGCTTGGGACGCTATGCAAGGGACCC
AAAATTTGTGTCAGCAACAAAACACGAAATCGCTGATGCCTGTGACCTCACCATCGACGA
GATGGAGAGTGCAGCCAGCACCCTGCTTAATGGGAACGTGCGTCCCCGAGCCAACGGGGA
TGTGGGCCCCCTCTCACACCGGCAGACTATGAGCTACAGGACTTTGGTCCTGGGCTTACA
GCGACGAAGAGCCAGACCCTGGGGAGGGATTGAGGGAGGACCTGGGCGGATGAATTGATT
TTGCNTCACCACCTTTGTTAGGCCCCCAGGCGAGGGGCAAG

R46658

TTTTTTTTTTNTTTTTTTTTTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTT
TAGAAAAATTTCTGTAGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTT
TCGCTGAATAAATGAGAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGNGTAAAAAT
ACTAAT

H29256

TTTTTTTTTTTTTTTTTTTTTGTGGAAAGATGATAGGTTTATAGNGACTCAAAATATTTT
AGAAAAATTTCTGTAGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTT
CGCTGAATAAATGAGAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATA
CTAATAATTTCTAGATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGC
AGTTCAAATACAATACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAANTTAT
AGNATATTACAAGTCATGTACAGTAAATCTATAATTTTGGACAANCTAGTGTATCTAAGT
TTACCNGGGGTGCGAGTGCCTTATTNTTCCNGTTTACAGTTGCCCTTAGCGTGACAGTCN
GGAACCGNCCTTC

H29339

GCCTGACTGTCCCCAGCAGCTTCCGGAACAAAAACAGCGACAAGCAGAGGAGTGCGGACA
NTTTGGTGGAGGCAGTCCTGATATCCGAAGCTTGGGACGCTATGCAAGGGACCCAAAATT
TGTGTCAGCAACAAAACACGAAATCGCTGATGCCTGTGACCTCACCATCGACGAGATGGA
GAGTGCAGCCAGCACCCTGCTTAATGGGAACGTGCGTCCCCGAGCCAACGGGGATGTGGG
CCCCCTCTCACACCGGCAGACTATGAGCTACAGGACTTTGGTCCTGGGCTACAGCGACGA
AGAGCCAGACCCTGGGAGGGATGAGGAGGAC

BG716371

AGCGGTCGTAATAATGTAGTTCCCCACTAAAATCTAGAAATTATTAGTATTTTTACTCGG
GCTATCCAGAAGTAGAAGAAATAGAGCAAATTCTCATTTATTCAGCGAAAATCCTCTGGG
GTTAAAATTTTAAGTTGAAAGAACTTGACACTACAGAAATTTTTCTAAAATATTTGAGTC
ACTATAAACCTATCATCTTTCCACAAGATATACCAGATGACTATTGCAGTCTTCTCTTGG
GCAAGAGTTCCATGATTTGATACTGTACCTTGGATCCACCATGGGTGCAACTGTCTTGGT
TTGTTGTTGACTTGAACCACCCTCTGGTAAGTAAGTGAATTACAGAGCAGGTCTAGCTGG
CTGCTCTGCCCCTTGGGTATCCATAGTTACGGTTTTCTCTGTGGCCCACCCAGGTGTTTT
TGCATCGCTGGTGCAGAAATGCACAGGTGGATGAGATATAGCTGCTCTTGTCCTCTGGGG
ACTGGTGGTGCTGCTTAAGAAATAAGGGGTGCTGGGGACAGAGGAGCAACGTGGTGATCT
ATAGGATTGGAGTGTCGGGGTCTGTACAAATCGTATTGTTGCCTTTTACAAAACTGTGTA
CTGTATGTTCTCTTTGAGGGCTTTTGTATGCAATTGAATGAGG

AI537488

TTTTCTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTTCTGT
AGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAAATGA

GAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTTCTAG
ATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAATACAAT
ACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGTATATTACAAGT
CATGTACAGTAAATCTATAATTTTAAACAAACCTGTGTATCTAAGTTTACCTGGTTG

AA458692

GACAAATAAAGCAATTATAAATGTATCTCACTTTAGAACAGACAAAAAAAGGGCATGCTA
TGGAAATTGTTTAAATCTCAAGCAACAATGCTGATTAATTTCTGGTCAATAATCGTTCTA
TAGTTCTCCTTCATGAAGCCTGGTGAGGTTCCAGGAAACAGCTTGATTTGGGAAGCCTCA
GCAGAAAAGAAAGCATCTCAGAGGACACATAAAATGTCTGGCAACCCCTCTTGGCGGCCC
TCATCCAGCAAAGCTTGTGTGGTCTTGGCAACTGTCCTCAGGACTCTGCTTTCAAGATGA
AAGAGGTGTAGCTTACCCGCTCAATACACCAAGTACAAGATTTAGTACGAAAAATGACCC
AAAGATGACGAGACTGACAAGATACACCCAGGGCAATTCCAATCCCATAGCATCATTCAT

AI393327

TTTATATTATTCACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATTCTTA
CTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAATCCT
TTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATACTTGG
GAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACAGGGCAGGTACTG
TGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGGTGCCAATG
GTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTACACCACTGTCAAC
ATTATCCTGGACTCTGTGTCTCTCTCTGTTGGGTCTTGTGGCATCACATCAGGCCAAAAT
TGCCAGACCAGGACCCTAAGTGTCTGATAGAGGCGATGATCTTTT

AI520947

TTTTTTTTTTTTTTTTTTTTTCTTACAAAGAAAAATTTAATATTCGATGAGAGGTTGAAC
CAGGCTTAAAGCAAACATACTAGGAAATGGGGCAGCCTGTAAGAATGCCAGTTTGTAAGT
ACTGACTTTGGAAAAGATCATCGCCTCTATCAGACACTTAGGGTCCTGGTCTGGCAATTT
TGGCCTGATGTGATGCCACAAGACCCAACAGAGAGAGACACAGAGTCCAGGATAATGTTG
ACAGGGGGTAGCCCTTTAGGAGAAATGGCGCTCCCTGCGGCTGGTATTAGGTTACCATTG
GCACCGAAGGAACCAGGAGGATAAGAATAT

AI248998

TGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATTCT
TACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAATC
CTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATACTT
GGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGGGCAGGTAC
TGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGGTGCCAA
TGGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTACACCACTGTCA
ACATTATCCTGGACTCTGTGTCTCTCTCTGTTGGGTCTTGTGGCATCACATCAGGCCAAA

ATTGCCAGACCAGGACCCTAAGTGTCTGATAGAGGCGATGATCTTTTCCAAAGTCAGTAC
TTACAAACTGGCATTCTTACAG

AI075844

TTTTTTTTTTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTT
AAGACATTCTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCT
GACATAAATCCTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAAT
TTTTATACTTGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACG
GGGCAGGTACTGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCT
TCGGTGCCAATGGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTAC
ACCACTGTCAACATTATCCTGGACTCTGTGTCTCTCTCTGTTGGGTCTTG

AI869807

GTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATTCTT
ACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAATCC
TTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATACTTG
GGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGGCAGGTACTG
TGCCAGGGCAGCTCTGAAATATGGATATTCTTACCTCCTGGTTCTTTCGGTGCAAATGGT
AACCTAATACCAGCCGCAGGGAGCGCCATTTCT

AI869800

GTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATTCTT
ACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAATCC
TTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATACTTG
GGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGNGCAGGTACT
GTGCCAGGGCAGCTCTGAATTATGGATATTCTTATCCTCCTG

AI243110

TTTTTCTTACAAAGAAAAATTTAATATTCGATGAGAGGTTGAACCAGGCTTAAAGCAGAC
ATACTAGGAAATGGTGCAGCCTGTAAGAATGCCAGTTTGTAAGTACTGACTTTGGAAAAG
ATCATCGCCTCTATCAGACACTTAGGGTCCTGGTCTGGCAATTTTGGCCTGATGTGATGC
CACAAGACCCAACAGAGAGAGACACAGAGTCCAGGATAATGTTGACAGTGGTGTAGCCCT
TTAGGAGAAATGGCGCTCCCTGCGGCTGGTATTAGGTTACCATTGGCACCGAAGGAACCA
GGAGGATAAGAATATCCATAATTTCAGAGCTGCCCTGGCACAGTACCTGCCCCGTCGGAG
GCTCTCACTGGCAAATGACAGCTCTGTGCAAGGAGCACTC

AI955764

TTATCTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTTCTGT
AGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAAATGA

GAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTTCTAG
ATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAATACAAT
ACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGGATATTACAAGG
CATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTTTACCTGGTTGCGA
GTGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGAAACCGACCATCGG
AGTGATATTCTCTTATGTAAACTGGCGTCACATCACAGAAAACCTTATTTATGAGGTCCC
AT

AA192669

GCCCTCACAGCCCACCACGCCTGGCCTTCGCCCAATTCTGAAACTTCGTAGGATAGAGCT
GGAAAGTGCCACATGGTGAAGCGAGATCCAGCTGTCTGGGTGGATGTCGGAGTCCATAGG
CTGAGCAGAGATGGTTCTTAGTGAGGTTCTCGCTGCCAGTTGACGGTGAAATCATAGCTG
CCATTTACATTTTGTGAGATTATGAAAAACATAAGACTAAAGAAACTAAATGTGTTATTC
CTGTGGACACAAAAATGTGTGTTTTTCAGATGGGGAGGGGACCAAAAAGGAAAAACATTT
CATCTTAAAACTTTCCTAAGACAAAGGAAAACAAAAAACCATGCTCCTACAACTTCAAAT
TTTTCTTACCAAAGAAAAATTTAATATTCGATGAGAGGTTGAACCAGGCTTAAAGCAGAC
ATACTAGGGAATGGGTGCAGCCTGTAAGAATGCCAGTTT

AA192157

GTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATTCTT
ACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAATCC
TTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATACTTG
GGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGGGCAGGTACT
GTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGGTGCCAAT
GGTAACCTAATACCAGCCGCAGGAGCGCCATTTCTCCTAAAGGGCTACACCACTGTCAAC
ATTATCCTGGGACTCTGTGTCTCTCTCTGTTGGGTCTTGTGGCATCACATCAGGCCAAAA
TTGGCCAGACCAGGACCCCAAGTGGTCTGATAGAAGGCGATGATCTTTTCCAAAGTCAGT
ACTTACA

AI361691

GTTTAAAATTATAGATTTACTGTACATGACTTGTAATATACTATAATTTGTATTTGTAAA
GAGATGGTCTATATTTTGTAATTACTGTATTGTATTTGAACTGCAGCAATATCCATGGGT
CCTAATAATTGTAGTTCCCCACTAAAATCTAGAAATTATTAGTATTTTTACTCGGGCTAT
CCAGAAGTAGAAGAAATAGAGCCAATTCTCATTTATTCAGCGAAAATCCTCTGGGGTTAA
AATTTTAAGTTTGAAAGAACTTGACACTACAGAAATTTTTCTAAAATATTTTGAGTCACT
ATAAACCTATCATCTTTCCACAAGATATACCAGATGACTATTTGCAGTCTTTTCTTTGGG
CAAGAGTTCCATGATTTTGATACTGTACCTTTGGATCCACCATGGGTTGCAACTGTCTTT
GGTTTTGTTTGTTTGACTTGAACCA

AI914244

TTCGCTGAATAAATGAGAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAA
TACTAATAATTTCTAGATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCT
GCAGTTCAAATACAATACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATT
ATAGTATATTACAAGTCATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAA
GTTTACCTGGTTGCGAGTGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTC
AGAAACCGACCAT

AW008769

TTTTATCTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTTCT
GTAGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAAAT
GAGAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTTCT
AGATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAATACA
ATACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGTATATTACAA
GTCATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTTTACCTGGTTGC
GAGTGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGAAACC

AW008794

TTTTATCTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTTCT
GTAGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAAAT
GAGAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTTCT
AGATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAATACA
ATACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGTATATTACAA
GTCATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTTTACCTGGTTGC
GAGTGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGAAACCGACCATC
GGAGTGATATTCTCTTATGTAAACTGGCGTCACATCACAGAAAACCTTATTTATTT

AA877582

TTTTTTTTTTAGAGCCAATTCTCATTTATTCAGCGAAAATCCTCTGGGGTTAAAATTTTA
AGTTTGAAAGAACTTGACACTACAGAAATTTTTCTAAAATATTTTGAGTCACTATAAACC
TATCATCTTTCCACAAGATATACCAGATGACTATTTGCAGTCTTTTCTTTGGGCAAGAGT
TCCATGATTTTGATACTGTACCTTTGGATCCACCATGGGTTGCAACTGTCTTTGGTTTTG
TTTGTTTGACTTGAACCACCCTCTGGTAAGTAAGTGAATTACAGAGCAGGTCCAGCTGGC
TGCTCTGCCCCTTGGGTATCCATAGTTACGGTTTTCTCTGTGGCCCACCCAGGGTGTTTT
TTGCATCGCTGGTGCAGAAATGCACAGGTGGATGAGATATAGCTGC

AI051972

TTTTTTTTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAA
GACATTCTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGA
CATAAATCCTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTT
TTATACTTGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACAGG

GCAGGTACTGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTC
GGTGCCAATGGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTACAC
CACTGTCAACATTATCCTGGACTCTGTGTCTCTCTCTGTTGAGTCTTGTGGCATCACATC
AGGCCAAAATTGCCAGACCAGGACCCTAAGTGTCTGATAGAGGCGATGATCTT

AI017959

TTTAGAGCCAATTCTCATTTATTCAGCGAAAATCCTCTGGGGTTAAAATTTTAAGTTTGA
AAGAACTTGACACTACAGAAATTTTTCTAAAATATTTTGAGTCACTATAAACCTATCATC
TTTCCACAAGATATACCAGATGACTATTTGCAGTCTTTTCTTTGGGCAAGAGTTCCATGA
TTTTGATACTGTACCTTTGGATCCACCATGGGTTGCAACTGTCTTTGGTTTTGTTTGTTT
GACTTGAACCACCCTCTGGTAAGTAAGTGAATTACAGAGCAGGTCCAGCTGGCTGCTCTG
CCCCTTGGGTATCCATAGTTACGGTTTTCTCTGTGGCCCACCCAGGGTGTTTTTTGCATC
GCTGGTGCAGAAATGCACAGGTGGATGAGATATAGCTGCTCTTGTCCTCTGGGGACTGGT
GGTGCTGCTTAAGAAATAAGGGGTGCTGGGGACAGAGGAGCAA

N79331

TTTTTTTTTTTTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTT
CTTAAGACATTCTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTT
CCTGACATAAATCCTTTTTG

N62240.1

ACAAAGAAAAATTTAATATTCGATGAGAGGTTGAACCAGGCTTAAAGCAGACATACTAGG
AAATGGTGCAGCCTGTAAGAATGCCAGTTTGTAAGTACTGACTTTGGAAAAGATCATCGC
CTCTATCAGACACTTAGGGTCCTGGTCTGGCAATTTTGGCCTGATGTGATGCCACAAGAC
CCAACAGAGAGAGACACAGAGTCCAGGNTAATATTGACAGNAGGTGGANGCCCCCCT

AI240933

TTTTTTTTTTTTTTTTTTTGGTCCAAAATTTTTAATAGTATACAGACAACCTGTTAATT
TTTTTTTTTTTTTTTTTTTGGAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACC
TCTTCTTAAGACATTCTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGGAAATGACT
CTTTCCTGACATAAATCCTTTTTTATTAAAATGCAAAAGGTTCTTCAGAATAAAACTGTG
TAATAATTTTTATACTTGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAG

AI015031

TTTTTCTTACAAAGAAAAATTTAATATTCGATGAGAGGTTGAACCAGGCTTAAAGCAGAC
ATACTAGGAAATGGTGCAGCCTGTAAGAATGCCAGTTTGTAAGTACTGACTTTGGAAAAG
ATCATCGCCTCTATCAGACACTTAGGGTCCTGGTCTGGCAATTTTGGCCTGATGTGATGC
CACAAGACCCAACAGAGAGAGACACAGAGTCCAGGATAATGTTGACAGTGGTGTAGCCCT
TTAGGAGAAATGGCGCTCCCTGCGGCTGGTATTAGGTTACCATTGGCACCGAAGAGACCA

GGAGGATAAGAATATCCATAATTTCAGAGCTGCCCTGGCACAGTACCTGCCCCGTCGGAG
GCTCTCACTGGCAAATGACAGCTCTGTGCAAGGAGCACTCCCAAGTATAAAAATTATTAC
ACAGTTTTATTCTG

AI290994

TAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATTCTTA
CTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAATCCT
TTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATACTTGG
GAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGGGCAGGTACTG
TGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGGTGCCAATG
GTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTACACCACTGTCAAC
ATTATCCTGGACTCTGTGTCTCTCTCTGTTGGGTCTTGTGGCATCACATCAGGCCAAAAT
TGCCAGACCAGGACCCTAAGTGTCTGATAGA

AA861160

TTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATT
CTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAA
TCCTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATAC
TTGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGAC

AA915941

TTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATTC
TTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAAT
CCTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATACT
TGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGAAGGGGCAGGTA
CTGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGGTGCCA
ATGGTAACCTAATACCAGCCGCAGGAGCGCCATTTCTCCTAAAGGGCTACACCACTGTCA
ACATTATCCTGGACTCTGTGTCTCTCTCTGTTGGGTCTTGTGGCATCACATCAGGCCAAA
ATTGCCAGACCAGGACCCTAAGTGTCTGATAGAGGCGATGATCTTTTCCAAAGTCAGTAC
TTA

AA493341

GCTCGACTTTTTTTTTGGGGGAACGTTTTCATTAGGTTAACAGTGTTTGGCAAGCATTGG
AAACACGGAATCTCACAGACAGATACAGGCAGAAAGAATCACAGTTCAATCCAAAAGCAA
CACACTGAGAGGACATCAGAGTCCAAACACATGCAGAGAAGCTGTCAGGGAGCAGCTAGG
AGACACGCAGAGTTGCCTCACACGTGGCAGCAGGAGAAGGTGCAACACGGATCCGACTGC
TTACCCACTAAGGACACCAAGAACCAGGTTAAGGACGAAAAATGAGCCAAGGATGATCAG
ACTAACAAAATACACCCATGGCCATTCCCATCCTATCGCATCATTTACCCAGTAGAGCAC
GTCTGTCCAGCCCTCCATGGTGATGCACTGAAACACAGTAAGCATGGCAAAGGCAAAGTT
ATCAAAGTTGGTGATGCCTCCGTTCGGGCCAACCCAGCCACTCCTACATTCCGTGCCATT

GGCAGTACACTGGCGTCCATTCCCTGT

AI467998

TTTTTTTTTTTTTTTTTGGTCCAAAATTTTTAATAGTATACAGACAACCTGTTAATTTTT
TTTTTTTTTTTTTTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCT
TCTTAAGACATTCTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTT
TCCTGACATAAATCCTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAA
TAATTTTTATACTTGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCC
GACGGGGCAGGTACTGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGT
TCCTTCGGTGCCAATGGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGG
CTACACCACTGTCAACATTATCC

AA885585

TTTTTTTTTTTTTTTTTTCTTACAAAGAAAAATTTAATATTCGATGAGAGGTTGAACCAG
GCTTAAAGCAGACATACTAGGAAATGGTGCAGCCTGTAAGAATGCCAGTTTGTAAGTACT
GACTTTGGAAAAGATCATCGCCTCTATCAGACACTTAGGGTCCTGGTCTGGCAATTTTGG
CCTGATGTGATGCCACAAGACCCAACAGAGAGAGACACAGAGTCCAGGATAATGTTGACA
GTGGTGTAGCCCTTTAGGAGAAATGGCGCTCCCTGCGGCTGGTATTAGGTTACCATTGGC
ACCGA

AI033648

TGTAAATAACAAACACCACTTGGTTATGAAGACCTTACAAACCTCTTCTTAAGACATTCT
TACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAATC
CTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATACTT
GGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGGGCAGGTAC
TGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGGTGCCAA
TGGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTACACCACTGTCA
ACATTATCCTGGACTC

AI697633

ATTCCTGTTAATTTTGACAAGCTCAACGGCTGAAATCTAGGAATGGTTACTACCAAAAGC
CCACCCAATCCAGCTCATTTTGCTATCGTTTTATAACAATTAATCTGCATTATATTTGGA
TCCAGACAAATAAAGCAATTATAAATGTATCTCACTTTACAACAGACAAAAAAAGGGCAT
GCTATGGAAATTGTTTAAATCTCAAGCAACAATGCTGATTAATTTCTGGTCAATAATCGT
TCTATAGTTCTCCTTCATGAAGCCTGGTGAGGTTCCAGGAAACAGCTTGATTTGGGAAGC
CTCAGCAGAAAAGAAAGCATCTCAGAGGACACATAAAATGTCTGGCAACCCCTCTTGGCG
GCCCTCATCCAGCAAAGCTTGTGTGGTCTTGGCAACTGTCCTCAGGACTCTGCTTTCAAG
ATGAAAGAGGTGTAGCTTACCCGCTCAATACACCAAGTACAAGATTTAGTACGAAAAATG
ACCCAAAGATGACGAGACTGACACAATACACCCAGGGCAATTCAAATCCCATAGCATCAT
TCAT

AA523647

GGTCGACGTATTTGTAAAGAGATGGTCTATATCTTGTAATTACTGTATTGTATTTGAACT
GCAGCAATATCCATGGGTCCTAATAATTGTAGTTCCCCACTAAAATCTAGAAATTATTAG
TATTTTTACTCGGGCTATCCAGAAGTAGAAGAAATAGAGCCAATTCTCATTTATTCAGCG
AAAATCCTCTGGGGTTAAAATTTTAAGTTTGAAAGAACTTGACACTACAGAAATTTTTCT
AAAATATTTTGAGTCACTATAAACCTATCATCTTTCCACAAGAAAAAAAAACAAAAAAAA
AGTCGACG

BQ710377

CAAAGTACTTCCCCACATTTAGCTGGATTTGTCTTTGGTTTGAAGAGGCTAATACGTGAA
AGATTTGTTCACAGTTGGATGTCCCCTTTTCTGAACCATGAAGTAATATTGTGAATGGAG
TTGAATGCTGAGGTTAGGGTGCCGGAAAGATTCAGGGTCCTTCGGTACCCTCACATGGCT
TGGCTTTGGTAGAACAAGAAACTAAGCTCTGATTTGGCTTTAAATGAGAGTGCTAAATTT
CCTTTTTCTAATAAAGAACCTAGCTAAACATTTATATATACTTTTGAACACTGAACTTTC
TTGTTGCAGAGTTAACAGCTGTTGGGGGTAGCTGACAGCTGGATCCTGGTGCTGTTGGTA
CCATGGTACCTGAAGTGCACAGGCTGGTAGCCACACCTGACATTAACAAGTGAGTGGTAA
CCTCTCTGCCGCTGGCTCACAGCTACTGTTTCCATAGAAATGGCTGTCGGGATCAGTGGA
AACGAGGTAAGTGAAAGTTTTCGCTGATCCTTGTTTCCATCAAGCTGACGTCTGTTTCCC
TGGCAACAGCAGTGGACAGCAGCCAGGCGCTAGCAACAGATTCAGTAGAGCTCTCACTTG
TCAGCTGTGGCTATCATCTGTTCCTGACCAAGTTCTTTTTTTTTTTTTTAATAATGTACA
GAAAGACCTCTGANGGACCAGGANGCNACTCTGGCCACATGTGCCCTCCTGGATGCTCGT
TTTGCAAATGGAGAGCTGTGTGCTGAGTTGACTTCTCTGTCCGCAGTTCCCCCTCCACTG
NGGCTCTGGGGTTGNTGATGTGCAGGTAAAAAAAAGGAGGGTTGTTGAAGGTTATTAGTT
GTTCCAAGGGGAAGCCTGTTGAAACCTGGTTGATCCCCAATCCCTATGGGGAAGAAAAAT
CTCTTTAAGGGGCTTTTCATGCCCAGAGACCCAAATTTT

BQ706920

GGTGGCGATTCGGACGAGGGCAAAGACTTCCCCCATTTAGCTGGATTTGTCTTTGGTTTG
AAGAGGCTAATACGTGAAAGATTTGTTCACAGTTGGATGTCCCCTTTTCTGAACCATGAA
GTAATATTGTGAATGGAGTTGAATGCTGAGGTTAGGGTGCCGGAAAGATTCAGGGTCCTT
CGGTACCCTCACATGGCTTGGCTTTGGTAGAACAAGAAACTAAGCTCTGATTTGGCTTTA
AATGAGAGTGCTAAATTTCCTTTTTCTAATAAAGAACCTAGCTAAACATTTATATATACT
TTTGAACACTGAACTTTCTTGTTGCAGAGTTAACAGCTGTTGGGGGTAGCTGACAGCTGG
ATCCTGGTGCTGTTGGTACCATGGTACCTGAAGTGCACAGGCTGGTAGCCACACCTGACA
TTAACAAGTGAGTGGTAACCTCTCTGCCGCTGGCTCACAGCTACTGTTTCCATAGAAATG
GCTGTCGGGATCAGTGGAAACGAGGTAAGTGAAAGTTTTCGCTGATCCTTGTTTCCATCA
AGCTGACGTCTGTTTCCCTGGCAACAGCAGTGGACAGCAGCCAGGCGCTAGCAACAGATT
CAGGAGAGCTCTCACTTGTCAGCTGTGGCTATCATCTGTTCCTGACCAAGTTCTTTTTTT
TTTTTTTAATAATGGACAGAAAGACCTCTGAGGACCCAGGAGGCACCTCTGGGCACATGT
GCCCTCCTGGATGCTCCTTTTGCAGATGGAGACCTGGGGGCTGAGTTGACTTCTCTGGCC

GCAGTTCCCCCTCCACCTGGGGCTCCTGGGTGGTGAGGGGCCAGGTAAAAAAAGGGAAGG
TGTTTGAGGGTATTAATGGGTCCCCGGGCGGGCTGATCGAATCCTGGGGACTCCACGTCC
CTGGGGGGACAAGAATCTCTTCAACGGGGTTTTCCGGCCGGGAGCCGGAGTTTTTTATTC
AGCGGG

BQ016847

TTTTTTTTTTTTTTTTTTCTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTT
AGAAAAATTTCTGTAGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTT
CGCTGAATAAATGAGAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATA
CTAATAATTTCTAGATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGC
AGTTCAAATACAATACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTAT
AGTATATTACAAGTCATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGT
TTACCTGGTTGCGAGTGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAG
AAACCGACCATCGGAGTGATATTCTCTTATGTAAACTGGCGTCACATCACAGAAAACCTT
ATTTATGAGGTCCCATTGCCCTCGCAATAATCACTGGTAGCTGGGTTCTGACTTACTTAC
ACACCGTATTTCAGAACAGCTAAACAGGAACCAGGACGCAGTGTATTTGGGGGAAAGGGT
TTACAAATGGATATGTTGGGCCCAGTGACTGATATGCTAGACCGATGGCTGAGGTAACGA
CACAGGTGTGATGATCGTCATCACCTTTAACT

CA943595

TGCAAATAAGGACAAGCTCAGCGGCTGAAATCTACAAATGGGGACTACCAAAAGCCCACC
CAATCCAGCTCATTTTGCTATCGTTTTATAACAATTAATCTGCATTATATTTGGATCCAG
ACAAATAAAGCAATTATAAATGTATCTCACTTTAGAACAGACAAAAAAAGGGCATGCTAT
GGAAATTGTTTAAATCTCAAGCAACAATGCTGATTAATTTCTGGTCAATAATCGTTCTAT
AGTTCTCCTTCATGAAGCCTGGTGAGGTTCCAGGAAACAGCTTGATTTGGGAAGCCTCAG
CAGAAAAGAAAGCATCTCAGAGGACACATAAAATGTCTGGCAACCCCTCTTGGCGGCCCT
CATCCAGCAAAGCTTGTGTGGTCTTGGCAACTGTCCTCAGGACTCTGCTTTCAAGATGAA
AGAGGTGTAGCTTACCCGCTCAATACACCAAGTACAAGATTTAGTACGAAAAATGACCCA
AAGATGACGAGACTGACAAAATACACCCAGGGCAATTCAAATCCCATAGCATCATTCATC
TGCAAGAAATAAGATGGTCTCATAGGAGTGGGTTAATAAGAGGATTTAATAAGGA

BM008196

GGCAAAGTACTTCCCCACATTTAGCTGGATTGGTCTTTGGTTTGAAGAGGCTAATACGTG
AAAGATTTGTTCACAGTTGGATGTCCCCTTTTCTGAACCATGAAGTAATATTGTGAATGG
AGTTGAATGCTGACGGTTAGGGTGCCGGAAAGATTCAGGGTCCTTCGGTACCCTCACATG
GCTTGGCTTTGGTAGAACAAGAAACTAAGCTCTGATTTGGCTTTAAATGAGAGTGCTAAA
TTTCCTTTTTCTAATAAAGAACCTAGCTAAACATTTATATATACTTTTGAACACTGAACT
TTCTTGTCAGCAGAGTTAACAGCTGTAGGGGGTAGCTGACACGGCTGGATCCTGGTGCTG
TTGGTACCATGGTACCTGAAGTGCACAGGCTGGTAGCCACACCTGACATTAACAACGTGA
GTGGTAACCTCTCTGCCGCTGGCTCACAGCTACTGTTTCCATCAGAAATGGCTGTCGGGC
TCACGTGGAAACGAGGTAAGTGAAAGTACGCTAGATCCTTGTTCCATCACAGCTGACGCT

CTGTTTCCCATGGCAACACCCAGCACGGACAAGCCGCCACGCCGCATAGACAACCACAAC
CACGTACAGCTCTCCACAAGTCAGCTCGTGGCTATCCATCATGTCCCTGAACAAGCCCAC
ACCACCCCCCCCCAAGCGACACAGCAACGAGCACCACCCGGACGAACCAAAGGACGGACC
CCCCTGCCCCAACCTCTCGCCCATCCGCGACAGACCCGCCAAGCAAACACGACAACCTAA
CAAAGCAGAGGGACAGACCCATAGCGCCCGCTACCGGAAGCGTACACCACTTCCCAACAG
TAAGGCCAAAAGAGCGACGCGGAGCACGTGAACGGATAAGAAAACGAGAGAAGGCACGGC
CGCATGGCAAACACACCAGCAAGCAGCAGACAGCACGTGGGCACGACACAGGACAGAAAG
CAGCCCACCTCAGAGGGGACCAACGAAGAGTCGCACGAC

BI769856

CTGGGCCCAACATATCCATTTTTAAACCCTTTCCCCCAAATACACTGCGTCCTGGTTCCT
GTTTAGCTGTTCTGAAATACGGTGTGTAAGTAAGTCAGAACCCAGCTACCAGTGATTATT
GCGAGGGCAATGGGACCTCATAAATAAGGTTTTCTGTGATGTGACGCCAGTTTACATAAG
AGAATATCACTCCGGTGGTCGGTTTCTGACTGTCACGCTAAGGGCAACTGTAAACTGGAA
TAATAATGCACTCGCAACCAGGTAAACTTAGATACACTAGTTTGTTTAAAATTATAGATT
TACTGTACATGACTTGTAATATACTATAATTTGTATTTGTAAAGAGATGGTCTATATTTT
GTAATTACTGTATTGTATTTGAACTGCAGCAATATCCATGGGTCCTAATAATTGTAGTTC
CCCACTAAAATCTAGAAATTATTAGTATTTTTACTCGGGCTATCCAGAAGTAGAAGAAAT
AGAGCCAATTCTCATTTATTCAGCGAAAATCCTCTGGGGTTAAAATTTTAAGTTTGAAAG
AACTTGACACTACAGAAATTTTTCTAAAATATTTTGAGTCACTATAAACCTATCATCTTT
CCACAAGATATACCAGATGACTATTTGCAGTCTTTTCTTTGGGCAAGAGTTCCATGATTT
TGATACTGTACCTTTGGATCCACCATGGGTTGCAN

BI758971

GGAAAAGAAATACTGTTTTAGAGAAATAACATTTTCAACAAAACATCCCTGGAGTCAGAT
TTTGAGTTGGGGTGGGCTAATCAGGGAGTCGGGGCTCTCTGCGTGATGTCAGTTCTATGG
CTAACTGGTTTTTCTAAACCAGCCAGCTGCCTATCAAAACAGTACAACTTTTCTAGGAAA
TGCAATTGGCAAAGACACTTACGATGCTGAGAAGTACACAAGGTGAAACTGCTCCAGTTT
TTCTCATAGCAGGGTCAGCAGGAAAGCAAGTGGTGCCCCTGGTCCCATCTCACACAGGTG
AGACTGCACCGAGAGGTAACGTGGCCCTCACAGCCCACCACGCCTGGCCTTCGCCCAATT
CTGAAACTTCGTAGGATAGAGCTGGAAAGTGCCACATGGTGAAGCGAGATCCAGCTGTCT
GGGTGGATGTCGGAGTCCATAGGCTGAGCAGAGATGGTTCTTAGTGAGGTTCTCGCTGCC
AGTTGACGGTGAAATCATAGCTGCCATTTACATTTTGTGAGATTATGAAAAACATAAGAC
TAAAGAAACTAAATGTGTTATTCCTGTGGACACAAAAATGTGTGTTTTTCAGATGGGGAG
GGGACCAAAAAGGAAAAACATTTCATCTTAAAACTTTCCTAAGACAAAGGAAAACAAAAA
ACCATGCTCTACAACTTCAAATTTTTCTTACAAAGAAAAATTTAATATTCGATGAGCAGG
TTGAACCAGGCTTAAAGCAGACATACTAGGAAATGGTGCAGCCTGTAAGAATGCCAGTTT
GTAAGTACTGACTTTGGAAAAGATCATCGCTCTATCAGACACTTAGGGTCCTGGTCTGGC
CATTTTGGCCTGATGTGATGCCAAAAGACC

AA468565

TTTTATCGTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTTCT
GTAGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAAAT
GAGAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTTCT
AGATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAATACA
ATACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGTATATTACAA
GTCATGTACAGTAAATCTATTTTAAACAAACTAGTGTATCTAAGTTTACCTGGTTGCGAG
TGCATTAT

AA437099

CTTACAAAGAAAAATTTAATATTCGATGAGAGGTTGAACCAGGCTTAAAGCAGACATACT
AGGAAATGGTGCAGCCTGTAAGAATGCCAGTTTGTAAGTACTGACTTTGGAAAAGATCAT
CGCCTCTATCAGACACTTAGGGTCCTGGTCTGGCAATTTTGGCCTGATGTGATGCCACAA
GACCCAACAGAGAGAGACACAGAGTCCAGGATAATGTTGACAGTGGTGTAGCCCTTTAGG
AGAAATGGCGCTCCCTGCGGCTGGTATTAGGTTACCATTGGCACCGAAGAGACCAGGAGG
ATAAGAATATCCATAATTTCAGAGCTGCCCTGGCACAGTACCTGCCCCGTCGGAGGCTCT
CACTGGCAAATGACAGCTCTGTGCAAGGAGCACTCCCAAGTATAAAAATTAT

CA867864

CCGCGTCCGGTCAGATGGTACAAGTTTGTCTCTATAATTAAGACTTTTCCACCATCACAA
ACTTTAAACACAAAGTCTAAAATCTTGGGCAGCATAGAAAATAGGTTCTAGCTAAGCAGG
AGTTTTGTCCTCTACCAAGACCTTTCCTGAAAATCACTTATCAAGACAGTTTCCTGTAAG
AAAAAGCCATATCCCAGCTGATTTTCCTTCCTGGGGCCAAAATCTGCTATTATTCGGCCT
GAAAGCCTTGATGACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
GTATGGATGCTTGTGTGTGTGTATGGGGAATATGTGATTAATGTGTGTTGGCTGCTGTTG
TCTCTGATTTGGCTACTGTTGTTTCTGATTTAAATCTAAGTAAATGTTTAATTAAATGTA
TAGAATGCTGTCTCTAATGTGACCCTCTCTCCTTATTAAATCCTCTTATTAACCCACTCC
TATGAGACCATCTTATTTCTTGCAGATGAATGATGCTATGGGATTTGAATTGCCCTGGGT
GTATTTTGTCAGTCTCGTCATCTTTGGGTCATTTTTCGTACTAAATCTTGTACTTGGTGT
ATTGAGCGGG

AA682690

AATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATACTTGGGATGTGCTCC
TTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCGACAGGCAGGTACTGTGCCAGGGCAG
CTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCTGTGCTCAATGGTAACCTAAT
ACCAGCCGCAGGACNCGCCATTTCTCCTAAAGGGCTACACCACTGTCAACATTATC

AA701888

TCAGCGAAAATCCTCTGGGGTTAAAATTTTAAGTTTGAAAGAACTTGACACTACAGAAAT
TTTTCTAAAATATTTTGAGTCACTATAAACCTATCATCTTTCCACAAGATATACCAGATG
ACTATTTGCAGTCTTTTCTTTGGGCAAGAGTTCCATGATTTTGATACTGTACCTTTGGAT

CCACCATGGGTTGCAACTGTCTTTGGTTTTGTTTGTTTGACTTGAACCACCCTCTGGTAA
GTAAGTAAGTGAATTACAGAGCAGGTCCAGCTGGCTGCTCTGCCCCTTGGGTATCCATAG
TTACGGTTTTCTCTGTGGCCCACCCAGGGTGTTTTTTGCATCGCTGGTGCAGAAATGCAT
AGGTGGATGAGATATAGCTGCTCTTGTCCTCTGGGGACTGGTGGTGCTGCTTAAGAAATA
AGGGGTG

BU182632

TTTTGTCAGTCTCGTCATCTTTGGGTCATTTTTCGTACTAAATCTTGTACTTGGTGTATT
GAGCGGGCACAGTGGCTCACGCCTATAATCCCAGCACTTTCGGAGGCCGAGGCAGCTGGA
CCACCCGAGATCAGGAGTTTGAGACCAGCCTGACTAAGGCAGTGAAACCCTGTCTCTACT
AAAAATACAAAAATTAGCCAGGCATGGTGGCGCATGCCTGTAATCCCAGCTACTTGGGAG
GCTGAGGCAGGAGAATCACTTGAACCAGGGAGGTGGAGATTGCAGTGAGCCAAGACTGCA
CCATTGCATTCCAGCCTGGGTGACAAGAGCAAAACTCCATCTCAAAAAAAAAAAAAAAAA
AAAAAAAAAAAGACTTTTCTCTCATTCAACACTTTACCAGCATCTACTGACAGAAAATGG
ACAATTGAATTTCCTCCAATATATATACCTCTGATATGTCTGCTTTGTAAAAGAGTAGTG
TAATTGCTTACAACATTGAAAAGGTTGTTATTGGGGTCCTGGGGTAGCCAGGATATCGGC
ATGATTTGTCACCATATTCAGAATAAAACTGTACTGCAATAGTGAGTTAATTCCATATCT
TGGCCAACAGAGAATTTTTGGCCAGTGGCTACTAAGGCACACGGAAGTCCAGTCTAAAAG
GGACAGGGGAGGACTCTTTGTAGATAGTTCTTATGATTAAAAAATAACTTCCTATGTGTT
GTAGTGATGATTTAAGCTGACAGAATGCTAAAGACACCCCTTATGATTACCTGGTAGCAA
AGTACCTTCCCCACATTTAACCTGGATTTGCCCTTTTGGGTTTGAAAGAGGCTAAATA

BQ898429

GGTGGGATTCGGCACGAGGGCAAGACTTCCCCACATTTAGCTGGATTTGTCTTTGGTTTG
AAGAGGCTAATACGTGAAAGATTTGTTCACAGTTGGATGTCCCCTTTTCTGAACCATGAA
GTAATATTTGTGATATGGAGTTCGAATGGCTGAGGTCTAGGTGTGCCGAGAAAGATTCAG
GGTCCTTCGGTACCCTCACATGGCTTGGCTTTGGTAGAACAAGAAACTAAGCTCTGATTT
GGCTTTAAATGAGAGTGCTAAATTTCCTTTTTCTAATAAAGAACCTAGCTAAACATTTAT
ATATACTTTTGAACACTGAACTTTCTTGTTGCAGAGTTAACAGCTGTTGGGGGTAGCTGA
CAGCTGGATCCTGGTGCTGTTGGTACCATGGTACCTGAAGTGCACAGGCTGGTAGCCACA
CCTGACATTAACAAGTGAGTGGTAACCTCTCTGCCGCTGGCTCACAGCTACTGTTTCCAT
AGAAATGGCTGTCGGGATCAGTGGAAACGAGGTAAGTGAAAGTTTTCGCTGATCCTTGTT
TCCATCAAGCTGACGTCTGTTTCCCTGGCAACAGCAGTGGACAGCAGCCAGGCGCTAGCA
ACAGATTCAGTAGAGCTCTCACTTGTCAGCTGTGGCTATCATCTGTTCCTGACCAAGTTC
TTTTTTTTTTTTTTAATAATGTACAGAAAGACCTCTGAGGACCCAGGAGGCACCTCTGGC
CACATGTGCCCTCCTGGATGCTCGTTTTGCAGATGGAGAGCTGTGTGCTGAGTTGACTTC
TCTGTCCGCAGTTCCCCCTCCACCTGTGCTCTGGGTTGTTGATGTGCCAGTTAAAACAGG
GAGGCTGCTTCAGGGTATTAGTGTTGCCAAGGGGAGGCTGTTGAAATCTGGTTGATCCCA
AATC

BQ711800

CAAAGTACTTCCCCACATTTAGCTGGATTTGTCTTTGGTTTGAAGAGGCTAATACGTGAA
AGATTTGTTCACAGTTGGATGTCCCCTTTTCTGAACCATGAAGTAATATTGTGAATGGAG
TTGAATGCTGAGGTTAGGGTGCCGGAAAGATTCAGGGTCCTTCGGTACCCTCACATGGCT
TGGCTTTGGTAGAACAAGAAACTAAGCTCTGATTTGGCTTTAAATGAGAGTGCTAAATTT
CCTTTTTCTAATAAAGAACCTAGCTAAACATTTATATATACTTTTGAACACTGAACTTTC
TTGTTGCAGAGTTAACAGCTGTTGGGGGTAGCTGACAGCTGGATCCTGGTGCTGTTGGTA
CCATGGTACCTGAAGTGCACAGGCTGGTAGCCACACCTGACATTAACAAGTGAGTGGTAA
CCTCTCTGCCGCTGGCTCACAGCTACTGTTTCCATAGAAATGGCTGTCGGGATCAGTGGA
AACGAGGTAAGTGAAAGTTTTCGCTGATCCTTGTTTCCATCAAGCTGACGTCTGTTTCCC
TGGCAACAGCAGTGGACAGCAGCCAGGCGCTAGCAACAGATTCAGTAGAGCTCTCACTTG
TCAGCTGTGGCTATCATCTGTTCCTGACCAAGTTCTTTTTTTTTTTTTTTAATAATGTACA
GAAAGACCTCTGAGGACCCAGGGAGCACCTCTGGCCACATGTGCCCTCCTGAATGCTCGT
TTTGCAAATGGAGAGCTGTGTGCTGAGTTGACTTCTCTGTCCGCAGGTCCCCCTCCAACT
GTGCTCCTGGGTTGTGATGTGCAGGGTTAAACCAGGGAAGCTGTTGAAGGGTATTAGTGT
TGCCAGGGAAAGGCTGTTGAATTCTGGTTGATCCCAAATCCCTAGGGGGAAGAGAAATCC
CTTACGAGTGGTTTTTCATGGCCAGGAACCCTATA

AA703120

TCAGCGAAAATCCTCTGGGGTTAAAATTTTAAGTTTGAAAGAACTTGACACTACAGAAAT
TTTTCTAAAATATTTTGAGTCACTATAAACCTATCATCTTTCCACAAGATATACCAGATG
ACTATTTGCAGTCTTTTCTTTGGGCAAGAGTTCCATGATTTTGATACTGTACCTTTGGAT
CCACCATGGGTTGCAACTGTCTTTGGTTTTGTTTGTTTGACTTGAACCACCCTCTGGTAA
GTAAGTAAGTGAATTACAGAGCAGGTCCAGCTGGCTGCTCTGCCCCTTGGGTATCCATAG
TTACGGTTTTCTCTGTGGCCCACCCAGGGTGTTTTTTGCATCGCTGGTGCAGAAATGCAT
AGGTGGATGAGATATAGCTGCT

AA978315

GTATATCTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTTCT
GTAGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAAAT
GAGAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTTCT
AGATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATAGCTGCAGTTCAAATACA
ATACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGTATATTACAA
GTCATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTTTACCAGGTTGC
GAGTGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGAAACCGACCATC
GGAGTGATATTCTCTTATGTAAACAGGCGTCACATCACAGA

BE550599

TTTTTTTTTTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTT
CTGTAGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAA
ATGAGAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTT
CTAGATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAATA

CAATACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGTATATTAC
AAGTCATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTTTACCTGGTT
GCGAGTGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGAAACCGACCA
TCGGAGTGATATTCTCTTATGTAAACTGGCGTCACATCACAGAAAACCTTATTTATGAGG
TCCCATTGCCCTCGCAATAATCACTGGTAGCTGGGTTCTGACTTACTTACACACCGTATT
TCAGAACAGCTAAACAG

BE502741

TTTGGTATATCTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAAT
TTCTGTAGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAAT
AAATGAGAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAAT
TTCTAGATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAA
TACAATACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGTATATT
ACAAGTCATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTTTACCTGG
TTGCGAGTGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGAAACCGAC
CATCGGAGTGATATTCTCTTATGTAAACTGGCGTCACATCACAGAAAACCTTATTTATGA
G

AW872382

TTTTTTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTTCTGT
AGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAAATGA
GAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTTCTAG
ATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAATACAAT
ACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGTATATTACAAGT
CATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTTTACCTGGTTGCGA
GTGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGAAACCGACCATCGG
AGTGATATTCTCTTATGTAAACTGGCGTCACATCACAGAAAACCTT

AW444663

CGGCCGCCAACTTTTTTGAATGAGTGAAGTGCCAGGTACCATGAGAAAACCCTAGCTGGT
AAAGATCAAACCTGAGTTAGTTCTAAATTCACATACGGATTTTTTTTGCATGACGAAATC
TATTCTCTTTTTCCTGACAACTTCTCCACCTAGATGTTTGGGAAAGTTGCCATGAGAGAT
AACAACCAGATCAATAGGAACAATAACTTCCAGACGTTTCCCCAGGCGGTGCTGCTGCTC
TTCAGGTGACTGCAACTGGCTTGGGCGGTGCTCCTGGGCAGGGGGGTCCGCTAGGCGTGG
GTCCAGAGGGACGGAGGACACAGGTTATTAAAGCAGTGTGCCTTTCTCAGTTG

AW341279

TAAATAACTAACACCATTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATTCTTA
CTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAATCCT
TTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATACTTG

GGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGGGCAGGTACT
GTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGGTGCCAAT
GGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTACACCACTGTCAA
CATTATCCTGGACTCTGTGTCTCTCTCTGTTGGGTCTTGTGGCATCACATCAGGCCAAAA
TTGCCAGACCAGGACCCTAAGTGTCTGATAGAGGCGATGATCTTTTCCAAAGTCAGTACT
TACAAACTGGCATTCTTACAGGCTGCACCATTTCCTAGTATGTCTG

CF456750

ACTTTTCTAGGAAATGCAATTGGCAAAGACACTTACGATGCTGAGAAGTACACAAGGTGA
AACTGCTCCAGTTTTTCTCATAGCAGGGTCAGCAGGAAAGCAAGTGGTGCCCCTGGTCCC
ATCTCACACAGGTGAGACTGCACCGAGAGGTAACGTGGCCCTCACAGCCCACCACGCCTG
GCCTTCGCCCAATTCTGAAACTTCGTAGGATAGAGCTGGAAAGTGCCACATGGTGAAGCG
AGATCCAGCTGTCTGGGTGGATGTCGGAGTCCATAGGCTGAGCAGAGATGGTTCTTAGTG
AGGTTCTCGCTGCCAGTTGACGGTGAAATCATAGCTGCCATTTACATTTTGTGAGATTAT
GAAAAACATAAGACTAAAGAAACTAAATGTGTTATTCCTGTGGACACAAAAATGTGTGTT
TTTCAGATGGGGAGGGGACCAAAAAGGAAAAACATTTCATCTTAAAACTTTCCTAAGACA
AAGGAAAACAAAAAACCATGCTCTACAACTTCAAATTTTTCTTACAAAGAAAAATTTAAT
ATTCGATGAGAGGTTGAACCAGGCTTAAAGCAGACATACTAGGAAATGGTGCAGCCTGTA
AGAATGCCAGTTTGTAAGTACTGACTTTGGAAAAGATCATCGCCTCTATCAGACACTTAG
GGTCCTGGTCTGGCAATTTTGGCCTGATGTGATGCCACAAGACCCAACAGAGAGAGACAC
AGAGTCCAGGATAATGTTGACAGTGGTGTA

AW139850

TTTTTTTTTTTTTTTTTAGAAGAAATAGAGCCAATTCTCATTTATTCAGCGAAAATCCTC
TGGGGTTAAAATTTTAAGTTTGAAAGAACTTGACACTACAGAAATTTTTCTAAAATATTT
TGAGTCACTATAAACCTATCATCTTTCCACAAGATATACCAGATGACTATTTGCAGTCTT
TTCTTTGGGCAAGAGTTCCATGATTTTGATACTGTACCTTTGGATCCACCATGGGTTGCA
ACTGTCTTTGGTTTTGTTTGTTTGACTTGAACCACCCTCTGGTAAGTAAGTGAATTACAG
AGCAGGTCCAGCTGGCTGCTCTGCCCCTTGGGTATCCATAGTTACGGTTTTCTCTGTGGC
CCACCCAGGGTGTTTTTTGCATCGCTGGTGCAGAAATGCACAGGTGGATGAGATATAGCT
GCTCTTGTCCTC

AW029633

TTATCTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTTCTGT
AGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAAATGA
GAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTTCTAG
ATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAATACAAT
ACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGTATATTACAAGT
CATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTTTACCTGGTTGCGA
GTGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGAAACCGACCATCGG
AGTGATATTCTCTTATGTAAACTGGCGTCACATCACAGAAAACCTTATTTATGAGGTCCC

ATTGCCCTCGCAATAATCACTG

AI963788

TTTTTCTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTTCTG
TAGTGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAAATG
AGAATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTTCTA
GATTTTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAATACAA
TACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGTATATTACAAG
TCATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTTTACCTGGT

AI951788

ATCTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTTCTGTAG
TGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAAATGAGA
ATTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTTCTAGAT
TTTAGTGGGGAACCTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAATACAATA
CAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGTATATTACAAGTC
ATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTTTACCTGGTTGCGAG
TGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGAAACCGACCATCGGA
GTGATATTCTCTTATGTAAACT

AI680744

TTTTCTTCAAATAATTACAAGCTCAGCGGCTGAAATCTACAAATGGGGACTACCAAAAGC
CCACCCAATCCAGCTCATTTTGCTATCGTTTTATAACAATTAATCTGCATTATATTTGGA
TCCAGACAAATAAAGCAATTATAAATGTATCTCACTTTAGAACAGACAAAAAAAGGGCAT
GCTATGGAAATTGTTTAAATCTCAAGCAACAATGCTGATTAATTTCTGGTCAATAATCGT
TCTATAGTTCTCCTTCATGAAGCCTGGTGAGGTTCCAGGGAAACAGCTTGATTTGGGAAG
CCTCAGCAGAAAAGAAAGCATCTCAGAGGACACATAAAATGTCTGGCAACCCCTCTTGGC
GGCCCTCATCCAGCAAAGCTTGTGTGGTCTTGGCAACTGTCCTCAGGACTCTGCTTTCAA
GATGAAAGAGGTGTAGCTTACCCGCTCAATACACCAAGTACAAGATTTAGTACGAAAAAT
GACCCAAAGATGACGAGACTGACAAAATACACCCAGGGCAATTCAAATCCCATAGCATCA
TTCATCTGCAAG

AI601252

TTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATT
CTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAA
TCCTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATAC
TTGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGGGCAGGT
ACTGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGGTGCC
AATGGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTACACCACTGT
CAACATTATCCTGGACTCTGTGTCTCTCTCTGTTGGGTCTTGTGGCATCACATCAGGCCA

AAATTGCCAGACCAGGACCCTAAGTGTCTGATAGAGGCGATGATCTTTTCCAAAGTCAGT
ACTTACAAACT

AI459166

TTTTTTTTTGGTCCAAAATTTTTAATAGTATACAGACAACCTGTTAATTTTTTTTTTTTT
TTTTTTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGA
CATTCTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACA
TAAATCCTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTT
ATACTTGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGGGC
AGGTACTGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGG
TGCCAATGGTAACCTAATACCAGCCGCAGGGAGCGCCATTT

AA885750

TCGACAGCTACCAGTGATTATTGCGAGGGCAATGGGACCTCATAAATAAGGTTTTCTGTG
ATGTGACGCCATTTACATAAGAGAATATCACTCCGATGGTCGGTTTCTGACTGTCACGCT
AAGGGCAACTGTAAACTGGAATAATAATGCACTCGCAACCAGGTAAACTTAGATACACTA
GTTTGTTTAAAATTATAGATTTACTGTACATGACTTGTAATATACTATAATTTGTATTTG
TAAAGAGATGGTCTATATTTTGTAATTACTGTATTGTATTTGAACTGCAGCAATATCCAT
GGGTCCTAATAATTGTAGTTCCCCACTAAAATCTAGAAATTATTAGTATTTTTACTCGGG
CTATCCAGAAGTAGAAGAAATAGAGCC

BX092736

GAATATGTGATTAATGTGTGTTGGCTGCTGTTGTCTCTGATTTGGCTACTGTTGTTTCTG
ATTTAAATCTAAGTAAATGTTTAATTAAATGTATAGAATGCTGTCTCTAATGTGACCCTC
TCTCCTTATTAAATCCTCTTATTAACCCACTCCTATGAGACCATCTTATTTCTTGCAGAT
GAATGATGCTATGGGATTTGAATTGCCCTGGGTGTATTTTGTCAGTCTCGTCATCTTTGG
GTCATTTTTCGTACTAAATCTTGTACTTGGTGTATTGAGCGGGTAAGCTACACCTCTTTC
ATCTTGAAAGCAGAGTCCTGAGGACAGTTGCCAAGACCACACAAGCTTTGCTGGATGAGG
GCCGCCAAGAGGGGTTGCCAGACATTTTATGTGTCCTCTGAGATGCTTTCTTTTCTGCTG
AGGCTTCCCAAATCAAGCTGTTTCCTGGAACCTCACCAGGCTTCATGAAGGAGA

BX114568

TTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATT
CTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAA
TCCTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATAC
TTGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGGGCAGGT
ACTGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGGTGCC
AATGGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTACACCACTGT
CAACATTATCCTGGACTCTGTGTCTCTCTCTGTTGGGTCTTGTGGCATCACATCAGGCCA
AAATTGCCAGACCAGGACCCTAAGTGTCTGATAGAGGCGATGATCTTTTCCAAAGTCAGT

ACTTACAAACTGGCATTCTTACAGGCTGCACCATTTCCTAGTATGTCTGCTTTAAGCCTG
GTTCAACCTCTCATCGAATATTAAATTTTTCTTTGTAAGAAAAAAAAAAAAAAA

BE672659

TTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATT
CTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAA
TCCTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATAC
TTGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACAGGGCAGGT
ACTGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGGTGCC
AATGGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTACACCACTGT
CAACATTATCCTGGACTCTGTGTCTCTCTCTGTTGGGTCTTGTGGCATCACATCAGGCCA
AAATTGCCAGACCAGGACCCTAAGTGTCTGATAGAGGCGATGATCTTTTCCAAAGTCAGT
ACTTACAAACTGGCATTCTTACAGGCTGCACCATTTCCTAGTATGTCTGCTTTAAGCCTG
GTTCAACC

N78509

GGAGAAAGGAGGGAAACCAGGAGCAGCCGGCATGGGCAGTGGCAGAATTGGCCCTGNTAG
AGAGCAGAGCTGATGCCATCCTTTTGGCAAATAGCTGACATTTTATGGTGTGGTGCTGGG
TGAGCCCCCTGTGAGGGTTGAACAGATGTGGACAGGACTTGGGTCCAGGCACTAGAGTGG
TGCAGCCTGTAAGAATGCCAGTTTGTAAGTACTGACTTTGGAAAAGATCATCGCCTCTAT
CAGACACTTAGGGTCCTGGTCTGGCAATTTTGGCCTGATGTGATGCCACAAGACCCAACA
GAGAGAGACACAGAGTCCAGGATNAATGTTGACAGTGGTGTAGCCTTTAGGAAGAAATGG
CGCTCCCTGCGGCTGGTATTAGGTTACCATTGGCANCCGAAGGAACCCAGGAGGATTAAG
AATTTCCCTAATTTCAGAACTTGCCCTGGCACAGTA

N73668

GGTCCAAAATTTTTAATAGTATACAGACAACCTGTTAATTTTTTTTTTTTTTTTTTTTGT
AAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATTCTTAC
TCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATCGACTCTTTCCTGACATAAATCCT
TTTTTATTAAAATNGCAAAATTGTTCTTCAGAATAAAACTGTGTAATAATTTTTATACTT
GGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGGGCAGGTAC
TGTGCCAGGGCAGCTCTGAAATTATGGAAATTCTTATCCCCCTGGTTCCTNCGGTGGCCA
ATGGGTAACCTAATACCAGCCCGCGGGAAGCGCCAATTTCNCCCAAAAGGGGGTAAACCA
CTGGTNAAACATTA

N46744

TTTTTCTTTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTA
AGACATTCTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTG
ACATAAATCCTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATT
TTTATANGTGGGGGNGCTC

N39597

ACAAAGAAAAATTTAATATTCGATGAGAGGTTGAACCAGGCTTAAAGCAGACATACTAGG
AAATGGTGCAGCCTGTAAGAATGCCAGTTTGTAAGTACTGACTTTGGAAAAGATCATCGC
CTCTATCAGACACTTAGGGTCCTGGTCTGGCAATTTTGGCCTGATGTGATGCCACAAGAC
CCAACAGAGAGAGACACAGAGTCCAGGATAATGTTGACAGTGGTGTAGCCCTTTAGGAGA
AATGGCGCTCCCTGCGGCTGGTATTAGGTTACCATTGGCACCGAAGAACCAGGAGGATAA
GAATATCCATAATTTCAGAGCTTGCCCTGGCACAGTACCTGCCCCGTCGGAGGCTCTCAC
TGGGCAAATGGACAGCTCTGTGCAAGGAGCACTCCCAAGTATAANAATTATTACACAGTT
TTATTCTGAAGAACATTTTGCATTTTAATAAAAAANGGA

BF439267

TTTTTTTTTTTTTTGGGCCAAAATTTTTAATAGTATACAGACAACCTGTTAATTTTTTTT
TTTTTTTTTTTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTTT
TAAGACATTCTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTTTTTCC
TGACATAAATCCTTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAA
TTTTTATACTTGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGAC
GGGGCAGGTACTGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCC
TTCGGTGCCAATGGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTA
CACCACTGTCAACATTATCCTGG

BF436153

TTTTTTTTTGGTCCAAAATTTTTAATAGTATACAGACAACCTGTTAATTTTTTTTTTTTT
TTTTTTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGA
CATTCTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACA
TAAATCCTTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTT
ATACTTGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGGGC
AGGTACTGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGG
TGCCAATGGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTACACCA
CTGTCAACATTATCCTGGACTC

BF110611

TTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTTCTGTAGTG
TCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAAAATGAGAA
TTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTTCTAGATT
TTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAATACAATACA
GTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGTATATTACAAGTCAT
GTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTTTACCTGGTTGCGAGTG
CATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGAAACCGACCATCGGAGT
GATATTCTCTTATGTAAACTGGCGTCACATCACAGAAAACCTTATTTATGA

M76558

gggcgagcgc ctccgtcccc ggatgtgagc tccggctgcc cgcggtcccg agccagcggc
gcgcgggcgg cggcggcggg caccgggcac cgcggcgggc gggcagacgg gcgggcatgg
ggggagcgcc gagcggcccc ggcggccggg ccggcatcac cgcggcgtct ctccgctaga
ggaggggaca agccagttct cctttgcagc aaaaaattac atgtatatat tattaagata
atatatacat tggattttat ttttttaaaa agtttatttt gctccatttt tgaaaaagag
agagcttggg tggcgagcgg ttttttttta aaatcaatta tccttatttt ctgttatttg
tccccgtccc tccccacccc cctgctgaag cgagaataag ggcagggacc gcggctccta
cctcttggtg atccccttcc ccattccgcc cccgccccaa cgcccagcac agtgccctgc
acacagtagt cgctcaataa atgttcgtgg atgatgatga tgatgatgat gaaaaaaatg
cagcatcaac ggcagcagca agcggaccac gcgaacgagg caaactatgc aagaggcacc
agacttcctc tttctggtga aggaccaact tctcagccga atagctccaa gcaaactgtc
ctgtcttggc aagctgcaat cgatgctgct agacaggcca aggctgccca aactatgagc
acctctgcac ccccacctgt aggatctctc tcccaaagaa aacgtcagca atacgccaag
agcaaaaaac agggtaactc gtccaacagc cgacctgccc gcgccctttt ctgtttatca
ctcaataacc ccatccgaag agcctgcatt agtatagtgg aatggaaacc atttgacata
tttatattat tggctatttt tgccaattgt gtggccttag ctatttacat cccattccct
gaagatgatt ctaattcaac aaatcataac ttggaaaaag tagaatatgc cttcctgatt
atttttacag tcgagacatt tttgaagatt atagcgtatg gattattgct acatcctaat
gcttatgtta ggaatggatg gaatttactg gattttgtta tagtaatagt aggattgttt
agtgtaattt tggaacaatt aaccaaagaa acagaaggcg ggaaccactc aagcggcaaa
tctggaggct ttgatgtcaa agccctccgt gcctttcgag tgttgcgacc acttcgacta
gtgtcaggag tgcccagttt acaagttgtc ctgaactcca ttataaaagc catggttccc
ctccttcaca tagccctttt ggtattattt gtaatcataa tctatgctat tataggattg
gaacttttta ttggaaaaat gcacaaaaca tgtttttttg ctgactcaga tatcgtagct
gaagaggacc cagctccatg tgcgttctca gggaatggac gccagtgtac tgccaatggc
acggaatgta ggagtggctg ggttggcccg aacggaggca tcaccaactt tgataacttt
gcctttgcca tgcttactgt gtttcagtgc atcaccatgg agggctggac agacgtgctc
tactggatga atgatgctat gggatttgaa ttgccctggg tgtattttgt cagtctcgtc
atctttgggt cattttcgt actaaatctt gtacttggtg tattgagcgg agaattctca
aaggaaagag agaaggcaaa agcacgggga gatttccaga agctccggga gaagcagcag
ctggaggagg atctaaaggg ctacttggat tggatcaccc aagctgagga catcgatccg
gagaatgagg aagaaggagg agaggaaggc aaacgaaata ctagcatgcc caccagcgag
actgagtctg tgaacacaga gaacgtcagc ggtgaaggcg agaaccgagg ctgctgtgga
agtctctgtc aagccatctc aaaatccaaa ctcagccgac gctggcgtcg ctggaaccga
ttcaatcgca gaagatgtag ggccgccgtg aagtctgtca cgttttactg gctggttatc
gtcctggtgt ttctgaacac cttaaccatt tcctctgagc actacaatca gccagattgg
ttgacacaga ttcaagatat tgccaacaaa gtcctcttgg ctctgttcac ctgcgagatg
ctggtaaaaa tgtacagctt gggcctccaa gcatatttcg tctctctttt caaccggttt
gattgcttcg tggtgtgtgg tggaatcact gagacgatct tggtggaact ggaaatcatg
tctcccctgg ggatctctgt gtttcggtgt gtgcgcctct taagaatctt caaagtgacc
aggcactgga cttccctgtg caacttagtg gcatccttat taaactccat gaagtccagt

gcttcgctgt tgcttctgct ttttctcttc attatcatct tttccttgct tgggatgcag
ctgtttggcg gcaagtttaa ttttgatgaa acgcaaacca agcggagcac ctttgacaat
ttccctcaag cacttctcac agtgttccag atcctgacag gcgaagactg gaatgctgtg
atgtacgatg gcatcatggc ttacgggggc ccatcctctt caggaatgat cgtctgcatc
tacttcatca tcctcttcat ttgtggtaac tatattctac tgaatgtctt cttggccatc
gctgtagaca atttggctga tgctgaaagt ctgaacactg ctcagaaaga agaagcggaa
gaaaaggaga ggaaaaagat tgccagaaaa gagagcctag aaaataaaaa gaacaacaaa
ccagaagtca accagatagc caacagtgac aacaaggtta caattgatga ctatagagaa
gaggatgaag acaaggaccc ctatccgcct tgcgatgtgc cagtagggga agaggaagag
gaagaggagg aggatgaacc tgaggttcct gccggacccc gtcctcgaag gatctcggag
ttgaacatga aggaaaaaat tgcccccatc cctgaaggga gcgctttctt cattcttagc
aagaccaacc cgatccgcgt aggctgccac aagctcatca accaccacat cttcaccaac
ctcatccttg tcttcatcat gctgagcagt gctgccctgg ccgcagagga ccccatccgc
agccactcct tccggaacac gatactgggt tactttgact atgccttcac agccatcttt
actgttgaga tcctgttgaa gatgacaact tttggagctt tcctccacaa aggggccttc
tgcaggaact acttcaattt gctggatatg ctggtggttg gggtgtctct ggtgtcattt
gggattcaat ccagtgccat ctccgttgtg aagattctga gggtcttaag ggtcctgcgt
cccctcaggg ccatcaacag agcaaaagga cttaagcacg tggtccagtg cgtcttcgtg
gccatccgga ccatcggcaa catcatgatc gtcaccaccc tcctgcagtt catgtttgcc
tgtatcgggg tccagttgtt caaggggaag ttctatcgct gtacggatga agccaaaagt
aaccctgaag aatgcagggg acttttcatc ctctacaagg atggggatgt tgacagtcct
gtggtccgtg aacggatctg gcaaaacagt gatttcaact tcgacaacgt cctctctgct
atgatggcgc tcttcacagt ctccacgttt gagggctggc ctgcgttgct gtataaagcc
atcgactcga atggagagaa catcggccca atctacaacc accgcgtgga gatctccatc
ttcttcatca tctacatcat cattgtagct ttcttcatga tgaacatctt tgtgggcttt
gtcatcgtta catttcagga acaaggagaa aaagagtata agaactgtga gctggacaaa
aatcagcgtc agtgtgttga atacgccttg aaagcacgtc ccttgcggag atacatcccc
aaaaacccct accagtacaa gttctggtac gtggtgaact cttcgccttt cgaatacatg
atgtttgtcc tcatcatgct caacacactc tgcttggcca tgcagcacta cgagcagtcc
aagatgttca atgatgccat ggacattctg aacatggtct tcaccggggt gttcaccgtc
gagatggttt tgaaagtcat cgcatttaag cctaaggggt attttagtga cgcctggaac
acgtttgact ccctcatcgt aatcggcagc attatagacg tggccctcag cgaagcagac
ccaactgaaa gtgaaaatgt ccctgtccca actgctacac ctgggaactc tgaagagagc
aatagaatct ccatcacctt tttccgtctt ttccgagtga tgcgattggt gaagcttctc
agcaggggag aaggcatccg gacattgctg tggactttta ttaagttctt tcaggcgctc
ccgtatgtgg ccctcctcat agccatgctg ttcttcatct atgcggtcat tggcatgcag
atgtttggga aagttgccat gagagataac aaccagatca ataggaacaa taacttccag
acgtttcccc aggcggtgct gctgctcttc aggtgtgcaa caggtgaggc ctggcaggag
atcatgctgg cctgtctccc agggaagctc tgtgaccctg agtcagatta caaccccggg
gaggagcata catgtgggag caactttgcc attgtctatt tcatcagttt ttacatgctc
tgtgcatttc tgatcatcaa tctgtttgtg gctgtcatca tggataattt cgactatctg
acccgggact ggtctatttt ggggcctcac catttagatg aattcaaaag aatatggtca
gaatatgacc ctgaggcaaa gggaaggata aaacaccttg atgtggtcac tctgcttcga
cgcatccagc ctcccctggg gtttgggaag ttatgtccac acagggtagc gtgcaagaga
ttagttgcca tgaacatgcc tctcaacagt gacgggacag tcatgtttaa tgcaaccctg

tttgctttgg ttcgaacggc tcttaagatc aagaccgaag ggaacctgga gcaagctaat
gaagaacttc gggctgtgat aaagaaaatt tggaagaaaa ccagcatgaa attacttgac
caagttgtcc ctccagctgg tgatgatgag gtaaccgtgg ggaagttcta tgccactttc
ctgatacagg actactttag gaaattcaag aaacggaaag aacaaggact ggtgggaaag
taccctgcga agaacaccac aattgcccta caggcgggat taaggacact gcatgacatt
gggccagaaa tccggcgtgc tatatcgtgt gatttgcaag atgacgagcc tgaggaaaca
aaacgagaag aagaagatga tgtgttcaaa agaaatggtg ccctgcttgg aaaccatgtc
aatcatgtta atagtgatag gagagattcc cttcagcaga ccaataccac ccaccgtccc
ctgcatgtcc aaaggccttc aattccacct gcaagtgata ctgagaaacc gctgtttcct
ccagcaggaa attcggtgtg tcataaccat cataaccata attccatagg aaagcaagtt
cccacctcaa caaatgccaa tctcaataat gccaatatgt ccaaagctgc ccatggaaag
cggcccagca ttgggaacct tgagcatgtg tctgaaaatg ggcatcattc ttcccacaag
catgaccggg agcctcagag aaggtccagt gtgaaaagaa cccgctatta tgaaacttac
attaggtccg actcaggaga tgaacagctc ccaactattt gccgggaaga cccagagata
catggctatt tcagggaccc ccactgcttg ggggagcagg agtatttcag tagtgaggaa
tgctacgagg atgacagctc gcccacctgg agcaggcaaa actatggcta ctacagcaga
tacccaggca gaaacatcga ctctgagagg ccccgaggct accatcatcc ccaaggattc
ttggaggacg atgactcgcc cgtttgctat gattcacgga gatctccaag gagacgccta
ctacctccca ccccagcatc ccaccggaga tcctccttca actttgagtg cctgcgccgg
cagagcagcc aggaagaggt cccgtcgtct cccatcttcc cccatcgcac ggccctgcct
ctgcatctaa tgcagcaaca gatcatggca gttgccggcc tagattcaag taaagcccag
aagtactcac cgagtcactc gacccggtcg tgggccaccc ctccagcaac ccctccctac
cgggactgga caccgtgcta cacccccctg atccaagtgg agcagtcaga ggccctggac
caggtgaacg gcagcctgcc gtccctgcac cgcagctcct ggtacacaga cgagcccgac
atctcctacc ggactttcac accagccagc ctgactgtcc ccagcagctt ccggaacaaa
aacagcgaca agcagaggag tgcggacagc ttggtggagg cagtcctgat atccgaaggc
ttgggacgct atgcaaggga cccaaaattt gtgtcagcaa caaaacacga aatcgctgat
gcctgtgacc tcaccatcga cgagatggag agtgcagcca gcaccctgct taatgggaac
gtgcgtcccc gagccaacgg ggatgtgggc cccctctcac accggcagga ctatgagcta
caggactttg gtcctggcta cagcgacgaa gagccagacc ctgggaggga tgaggaggac
ctggcggatg aaatgatatg catcaccacc ttgtagcccc cagcgagggg cagactggct
ctggcctcag gtggggcgca ggagagccag gggaaaagtg cctcatagtt aggaaagttt
aggcactagt tgggagtaat attcaattaa ttagactttt gtataagaga tgtcatgcct
caagaaagcc ataaacctgg taggaacagg tcccaagcgg ttgagcctgg cagagtacca
tgcgctcggc cccagctgca ggaaacagca ggccccgccc tctcacagag gatgggtgag
gaggccagac ctgccctgcc ccattgtcca gatgggcact gctgtggagt ctgcttctcc
catgtaccag ggcaccaggc ccacccaact gaaggcatgg cggcggggtg caggggaaag
ttaaaggtga tgacgatcat cacacctgtg tcgttacctc agccatcggt ctagcatatc
agtcactggg cccaacatat ccatttttaa acccttccc ccaaatacac tgcgtcctgg
ttcctgttta gctgttctga aatacggtgt gtaagtaagt cagaacccag ctaccagtga
ttattgcgag ggcaatggga cctcataaat aaggttttct gtgatgtgac gccagtttac
ataagagaat atcac

AF088004

tttttttttt cttacaaaga aaaatttaat attcgatgag aggttgaacc aggcttaaag
cagacatact aggaaatggt gcagcctgta agaatgccag tttgtaagta ctgactttgg
aaaagatcat cgcctctatc agacacttag ggtcctggtc tggcaatttt ggcctgatgt
gatgccacaa gacccaacag agagagacac agagtccagg ataatgttga cagtggtgta
gccctttagg agaaatggcg ctccctgcgg ctggtattag gttaccattg gcaccgaagg
aaccaggagg ataagaatat ccataatttc agagctgccc tggcacagta cctgccccgt
cggaggctct cactggcaaa tgacagctct gtgcaaggag cactcccaag tataaaaatt
attacacagt tttattctga agaacatttt gcattttaat aaaaaaggat ttatgtcagg
aaagagtcat ttacaaacct tgaagtgttt ttgcctggat cagagtaaga atgtcttaag
aagaggtttg taaggtcttc ataacaaagt ggtgtttgtt atttacaaaa aaaaaaaaaa
aaaaaaatta acaggttgtc tgtatactat taaaaat

M83566

agaataaggg cagggaccgc ggctcctatc tcttggtgat ccccttcccc attccgcccc
cgcctcaacg cccagcacag tgccctgcac acagtagtcg ctcaataaat gttcgtggat
gatgatgatg atgatgatga aaaaaatgca gcatcaacgg cagcagcaag cggaccacgc
gaacgaggca aactatgcaa gaggcaccag acttcctctt tctggtgaag gaccaacttc
tcagccgaat agctccaagc aaactgtcct gtcttggcaa gctgcaatcg atgctgctag
acaggccaag gctgcccaaa ctatgagcac ctctgcaccc ccacctgtag gatctctctc
ccaaagaaaa cgtcagcaat acgccaagag caaaaaacag ggtaactcgt ccaacagccg
acctgcccgc gccctttttct gtttatcact caataacccc atccgaagag cctgcattag
tatagtggaa tggaaaccat ttgacatatt tatattattg gctatttttg ccaattgtgt
ggccttagct atttacatcc cattccctga agatgattct aattcaacaa atcataactt
ggaaaaagta gaatatgcct tcctgattat ttttacagtc gagacatttt tgaagattat
agcgtatgga ttattgctac atcctaatgc ttatgttagg aatggatgga atttactgga
ttttgttata gtaatagtag gattgtttag tgtaattttg gaacaattaa ccaaagaaac
agaaggcggg aaccactcaa gcggcaaatc tggaggcttt gatgtcaaag ccctccgtgc
ctttcgagtg ttgcgaccac ttcgactagt gtcaggggtg cccagtttac aagttgtcct
gaactccatt ataaaagcca tggttcccct ccttcacata gccctttttgg tattatttgt
aatcataatc tatgctatta taggattgga actttttatt ggaaaaatgc acaaaacatg
ttttttttgct gactcagata tcgtagctga agaggaccca gctccatgtg cgttctcagg
gaatggacgc cagtgtactg ccaatggcac ggaatgtagg agtggctggg ttggcccgaa
cggaggcatc accaactttg ataactttgc ctttgccatg cttactgtgt ttcagtgcat
caccatggag ggctggacag acgtgctcta ctgggtaaat gatgcgatag gatgggaatg
gccatgggtg tattttgtta gtctgatcat ccttggctca tttttcgtcc ttaacctggt
tcttggtgtc cttagtggag aattctcaaa ggaaagagag aaggcaaaag cacggggaga
tttccagaag ctccgggaga agcagcagct ggaggaggat ctaaagggct acttggattg
gatcacccaa gctgaggaca tcgatccgga gaatgaggaa gaaggaggag aggaaggcaa
acgaaatact agcatgccca ccagcgagac tgagtctgtg aacacagaga acgtcagcgg
tgaaggcgag aaccgaggct gctgtggaag tctctggtgc tggtggagac ggagaggcgc
ggccaaggcg gggccctctg ggtgtcggcg gtggggtcaa gccatctcaa aatccaaact
cagccgacgc tggcgtcgct ggaaccgatt caatcgcaga agatgtaggg ccgccgtgaa
gtctgtcacg ttttactggc tggttatcgt cctggtgttt ctgaacacct taaccatttc

ctctgagcac tacaatcagc cagattggtt gacacagatt caagatattg ccaacaaagt
cctcttggct ctgttcacct gcgagatgct ggtaaaaatg tacagcttgg gcctccaagc
atatttcgtc tctcttttca accggtttga ttgcttcgtg gtgtgtggtg gaatcactga
gacgatcctg gtggaactgg aaatcatgtc tcccctgggg atctctgtgt ttcggtgtgt
gcgcctctta agaatcttca aagtgaccag gcactggact tccctgagca acttagtggc
atccttatta aactccatga agtccatcgc ttcgctgttg cttctgcttt ttctcttcat
tatcatcttt tccttgcttg ggatgcagct gtttggcggc aagtttaatt ttgatgaaac
gcaaaccaag cggagcacct ttgacaattt ccctcaagca cttctcacag tgttccagat
cctgacaggc gaagactgga atgctgtgat gtacgatggc atcatggctt acgggggccc
atcctcttca ggaatgatcg tctgcatcta cttcatcatc ctcttcattt gtggtaacta
tattctactg aatgtcttct tggccatcgc tgtagacaat ttggctgatg ctgaaagtct
gaacactgct cagaaagaag aagcggaaga aaaggagagg aaaaagattg ccagaaaaga
gagcctagaa aataaaaaga acaacaaacc agaagtcaac cagatagcca acagtgacaa
caaggttaca attgatgact atagagaaga ggatgaagac aaggacccct atccgccttg
cgatgtgcca gtaggggaag aggaagagga agaggaggag gatgaacctg aggttcctgc
cggaccccgt cctcgaagga tctcggagtt gaacatgaag gaaaaaattg cccccatccc
tgaagggagc gctttcttca ttcttagcaa gaccaacccg atccgcgtag gctgccacaa
gctcatcaac caccacatct tcaccaacct catccttgtc ttcatcatgc tgagcagcgc
tgccctggcc gcagaggacc ccatccgcag ccactccttc cggaacacga tactgggtta
ctttgactat gccttcacag ccatctttac tgttgagatc ctgttgaaga tgacaacttt
tggagctttc ctccacaaag gggccttctg caggaactac ttcaatttgc tggatatgct
ggtggttggg gtgtctctgg tgtcatttgg gattcaatcc agtgccatct ccgttgtgaa
gattctgagg gtcttaaggg tcctgcgtcc cctcagggcc atcaacagag caaaaggact
taagcacgtg gtccagtgcg tcttcgtggc catccggacc atcggcaaca tcatgatcgt
cactaccctc ctgcagttca tgtttgcctg tatcggggtc cagttgttca aggggaagtt
ctatcgctgt acggatgaag ccaaaagtaa ccctgaagaa tgcaggggac ttttcatcct
ctacaaggat ggggatgttg acagtcctgt ggtccgtgaa cggatctggc aaaacagtga
tttcaacttc gacaacgtcc tctctgctat gatggcgctc ttcacagtct ccacgtttga
gggctggcct gcgttgctgt ataaagccat cgactcgaat ggagagaaca tcggcccaat
ctacaaccac cgcgtggaga tctccatctt cttcatcatc tacatcatca ttgtagcttt
cttcatgatg aacatctttg tgggctttgt catcgttaca tttcaggaac aaggagaaaa
agagtataag aactgtgagc tggacaaaaa tcagcgtcag tgtgttgaat acgccttgaa
agcacgtccc ttgcggagat acatccccaa aaacccctac cagtacaagt tctggtacgt
ggtgaactct tcgcctttcg aatacatgat gtttgtcctc atcatgctca acacactctg
cttggccatg cagcactacg agcagtccaa gatgttcaat gatgccatgg acattctgaa
catggtcttc accggggtgt tcaccgtcga gatggttttg aaagtcatcg catttaagcc
taaggggtat tttagtgacg cctggaacac gtttgactcc ctcatcgtaa tcggcagcat
tatagacgtg gccctcagcg aagcggaccc aactgaaagt gaaaatgtcc ctgtcccaac
tgctacacct gggaactctg aagagagcaa tagaatctcc atcacctttt tccgtctttt
ccgagtgatg cgattggtga agcttctcag caggggggaa ggcatccgga cattgctgtg
gacttttatt aagtcctttc aggcgctccc gtatgtggcc ctcctcatag ccatgctgtt
cttcatctat gcggtcattg gcatgcagat gtttgggaaa gttgccatga gagataacaa
ccagatcaat aggaacaata acttccagac gtttccccag gcggtgctgc tgctcttcag
gtgtgcaaca ggtgaggcct ggcaggagat catgctggcc tgtctcccag ggaagctctg
tgaccctgag tcagattaca accccgggga ggagtataca tgtgggagca actttgccat

tgtctatttc atcagttttt acatgctctg tgcatttctg atcatcaatc tgtttgtggc
tgtcatcatg gataatttcg actatctgac ccgggactgg tctattttgg ggcctcacca
tttagatgaa ttcaaaagaa tatggtcaga atatgaccct gaggcaaagg gaaggataaa
acaccttgat gtggtcactc tgcttcgacg catccagcct cccctggggt ttgggaagtt
atgtccacac agggtagcgt gcaagagatt agttgccatg aacatgcctc tcaacagtga
cgggacagtc atgtttaatg caaccctgtt tgctttggtt cgaacggctc ttaagatcaa
gaccgaaggg aacctggagc aagctaatga agaacttcgg gctgtgataa agaaaatttg
gaagaaaacc agcatgaaat tacttgacca agttgtccct ccagctggtg atgatgaggt
aaccgtgggg aagttctatg ccactttcct gatacaggac tactttagga aattcaagaa
acggaaagaa caaggactgg tgggaaagta ccctgcgaag aacaccacaa ttgccctaca
ggcgggatta aggacactgc atgacattgg gccagaaatc cggcgtgcta tatcgtgtga
tttgcaagat gacgagcctg aggaaacaaa acgagaagaa gaagatgatg tgttcaaaag
aaatggtgcc ctgcttggaa accatgtcaa tcatgttaat agtgatagga gagattccct
tcagcagacc aataccaccc accgtcccct gcatgtccaa aggccttcaa ttccacctgc
aagtgatact gagaaaccgc tgtttcctcc agcaggaaat tcggtgtgtc ataaccatca
taaccataat tccataggaa agcaagttcc cacctcaaca aatgccaatc tcaataatgc
caatatgtcc aaagctgccc atggaaagcg gcccagcatt gggaaccttg agcatgtgtc
tgaaaatggg catcattctt cccacaagca tgaccgggag cctcagagaa ggtccagtgt
gaaaagaacc cgctattatg aaacttacat taggtccgac tcaggagatg aacagctccc
aactatttgc cgggaagacc cagagataca tggctatttc agggaccccc actgcttggg
ggagcaggag tatttcagta gtgaggaatg ctacgaggat gacagctcgc ccacctggag
caggcaaaac tatggctact acagcagata cccaggcaga aacatcgact ctgagaggcc
ccgaggctac catcatcccc aaggattctt ggaggacgat gactcgcccg tttgctatga
ttcacggaga tctccaagga gacgcctact acctcccacc ccagcatccc accggagatc
ctccttcaac tttgagtgcc tgcgccggca gagcagccag gaagaggtcc cgtcgtctcc
catcttcccc catcgcacgg ccctgcctct gcatctaatg cagcaacaga tcatggcagt
tgccggccta gattcaagta aagcccagaa gtactcaccg agtcactcga cccggtcgtg
ggccaccccc ccagcaaccc ctccctaccg ggactggaca ccgtgctaca ccccccctgat
ccaagtggag cagtcagagg ccctggacca ggtgaacggc agcctgccgt ccctgcaccg
cagctcctgg tacacagacg agcccgacat ctcctaccgg actttcacac cagccagcct
gactgtcccc agcagcttcc ggaacaaaaa cagcgacaag cagaggagtg cggacagctt
ggtggaggca gtcctgatat ccgaaggctt gggacgctat gcaagggacc caaaatttgt
gtcagcaaca aaacacgaaa tcgctgatgc ctgtgacctc accatcgacg agatggagag
tgcagccagc accctgctta atgggaacgt gcgtccccga gccaacgggg atgtgggccc
cctctcacac cggcaggact atgagctaca ggactttggt cctggctaca gcgacgaaga
gccagaccct gggagggatg aggaggacct ggcggatgaa atgatatgca tcaccacctt
gtagccccca gcgaggggca gactggctct ggcctcaggt ggggcgcagg agagccaggg
gaaaagtgcc tcatagttag gaaagtttag gcactagttg ggagtaatat tcaattaatt
agacttttgt ataagagatg tcatgcctca agaaagccat aaacctggta ggaacaggtc
ccaagcggtt gagcctggca gagtaccatg cgctcggccc cagctgcagg aaacagcagg
ccccgccctc tcacagagga tgggtgagga ggccagacct gccctgcccc attgtccaga
tgggcactgc tgtggagtct gcttctccca tgtaccaggg caccaggccc acccaactga
aggcatggcg gcggggtgca ggggaaagtt aaaggtgatg acgatcatca cacctcgtgt
cgttacctca gccatcggtc tagcatatca gtcactgggc ccaacatatc catttttaaa
ccctttcccc caaatacact gcgtcctggt tcctgtttag ctgttctgaa ata

CB410657

GTACTGTGCCGGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGGTG
CCAATGGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTACACCACT
GTCAACATTATCCTGGACTCTGTGTCTCTCTCTGTTGGGTCTTGTGGCATCACATCAGGC
CAAAATTGCCAGACCAGGACCCTAAGTGTCTGATAGAGGCGATGATCTTTTCAAAGTCAG
TAC

BQ372430

TGCAGCAANTGGCACGGAATGTAGGAGTGGGTGGGTGGGACCGAACGGAGGCATCACCAA
CTTTGATAACTTGGCCTATGCCATGCTTACGGTGTTTCAGTGCATCACCATGGAGGGCTG
GACAGATGTGCTCTACTGGGTAAATGATGCGATAGGATGGGAATGGCCATGGGCGTATTT
TGTTAGTCTGATCATCCTTGGCTCATTTTTCGTCCTTAACCTGGTTCTTGGTGTCCTTAG
TGGAGAATTCTCAAAGGAAAGAGAGAAGGCAAAAGCACGGGGAGATTTCCAGAAGCTCCG
GGAGAAGCAGCAGCTGGAGGAGGATCTAAAGGGCTACTTGG

BQ366601

ATGACTACGGGGGAAGTTCATTCTGACCTTCCAGACTAGCTAGTACTATATGAAATCCGA
GAGACGGAATGAACACGGACTGATGGGAAAGTACCCTGCGAAGAACACCACAATTGCCCT
ACAGGCGTGATTAAGGACACTGCATGATAGTTGCTCCAGAATGCCGGCGTGCTATATCGT
GTGATTTGCAAGATGACGAGCGTGAGGAAACAAAACGAGAAGAAGAAGATGATGTGTTCA
AAAGAAATGGTGCCCTGCTTGGAAACCATGTCAATCATGTTAATAGTGATAGGAGAGATT
CCCTTCAGCAGACCAATACCACCCACCGTCCNCTGCATGTCCAAAGGCCTTCAATTCCAC
CTGCAAGTGATACTGAGAAACCGCTGTTCCTCCAGCAGGAAATTCG

BQ324528

TACATCTCCGCTATCTGTGCCGTGTAACACGGTGTCCAGTCTCGTTAGGGAGGGGCTGCT
GGAGGGGTGGCCCACGACCGGGTCGAGTGACTCGGTGAGCACTTCTGTGCTTTACTTGAA
TCTAGGCCGGCAACTGCCATGATCTGTTGCTGCATTAGATGCAGAGGCAGTGCCGCGCGA
TGGTGAAGATGGGAGACGACGGGACCTCTTGCTGGCTGCTCTGCCGGCGCAGGCAC

BQ318830

TGTCGTGACTGGCGATACCTGGCGTTAGTGTGTACATGGTGTTCATAATTGCTGCTGCAT
AACATTTTGTGAGAATTAATGTGACAATGTATGTGCAGTGCTTAGCACATAGCAAGTGCT
CATGAATGGTAGCCACCAAGATGGCTGTTGTCATTTTAGTTTGCAGCAGTTCCACTTGTC
ATCATTGAGTTCCCAGGGAGTCCCCTCTTCTTTGGGAACAGACTTGCTCTCTGTAGCTCC
ATTGCGGTAAAAACAGATGAGGTTAATCCCTGTCCCAATCATTTTGGAGATGGCGTCGTT
TGTATTCCAATTCCACAGCCCAGTTCTTGTCTTTGTCTTCCTTTTATTTAAGCAGCAGCC
ACACAGAATTAGCCCTTTTCAAAAATAAATAAGATTATCATCCTGTTTTGCGTCCCTGGG

GTAACAGACTCTAACATTTCTTTCTCTTTCTCTTCTTTCAGATTGTCTAGTGTAATTTTG
GAACAATTAACCAAAGAAACAGAAGGCGGGAACCACTCACGCGGCAAATCTGGAGGCTTT
GATGTCAAAGCCCTCCGTGCCTTTCGAGTGTTGCGACCACTTCGAA

AL708030

AGTTCCCACCTCAACAAATGCCAATCTCAATAATGCCAATATGTCCAAAGCTGCCCATGG
AAAGCGGCCCAGCATTGGGAACCTTGAGCATGTGTCTGAAAATGGGCATCATTCTTCCCA
CAAGCATGACCGGGAGCCTCAGAGAAGGTCCAGTGTGAAAAGGTCCGACTCAGGAGATGA
ACAGCTCCCAACTATTTGCCGGGAAGACCCAGAGATACATGGCTATTTCAGGGACCCCCA
CTGCTTGGGGGAGCAGGAGTATTTCAGTAGTGAGGAATGCTACGAGGATGACAGCTCGCC
CACCTGGAGCAGGCAAAACTATGGCTACTACAGCAGATACCCAGGCAGAAACATCGACTC
TGAGAGGCCCCGAGGCTACCATCATCCCCAAGGATTCTTGGAGGACGATGACTCGCCCGT
TTGCTATGATTCACGGAGATCTCCAAGGAGACGCCTACTACCTCCCACCCCAGCATGTGA
GGCCAGATTTTTTGTTTTTGGGTGGAACCTCCCGGGGAACAGTGTACCTTTCCCCCAACC
CCCGCTCTG

BM509161

ATTCGGCACGAGCCTCCTTCAACTTTGAGTGCTCTGCCCCTTGGGTATCCATAGTTACGG
TTTTCTCTGTGGCCCACCCAGGGTGTTTTTTGCATCGCTGGTGCAGAAATGCACAGGTGG
ATGAGATATAGCTGCTCTTGTCCTCTGGGGACTGGTGGTGCTGCTTAAGAAATAAGGGGT
GCTGGGGACAGAGGAGCAACGTGGTGATCTATAGGATTGGAGTGTCGGGGTCTGTACAAA
TCGTATTGTTGCCTTTTACAAAACTGCTGTACTGTATGTTCTCTTTGAGGGCTTTTATAT
GCAATTGACTGAGGGCTGAAGTTTTCATTAGAATGCACTCACACTCTGACTGTACGTCCT
GATGAAAACCCACTTTTGGATAATTAGAACCGTCAAGGCTTCATTTTCTGTCAACAGAAT
TAGGCCGACTGTCAGGTTACCTTGGCAGGGATTCCCTGCAATCAAAAAGATAGATGATAG
GTAGCAATTTTGGTCCAAAATTTTTAATAGTATACAGACAACCTGTTAATTTTTTTTTTT
TTTTTTTTTGTAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTT

N85902

GGAAAACTCAAGTCCAGAGCAATACTACGTAAAATTCAGAAGTGAGAACATACAAAGGCA
ACACACAGGCTGACGAAGAAACAGAAAGAAGATACTGACCTGAGTTTGGATTTTGAGATG
GCTTGACTGAAAGAAAGACAAAAAGTGTTAAGATTCTGGTTCCGAGGGCTTGAGCACACA
CTCCCCATCATTTCAGCTGGAGATTTCAT

BQ774355

TTTTTTTTTTTTTTTTTTATTCTGAAGAACATTTTGCATTTTAATAAAAAAGGATTTATG
TCAGGAAAGAGTCATTTACAAACCTTGAAGTGTTTTTGCCTGGATCAGAGTAAGAATGTC
TTAAGAAGAGGTTTGTAAGGTCTTCATAACAAAGTGGTGTTTGTTATTTACAAAAAAAAA
AAAAAAAAATTAACAGGTTGTCTGTATACTATTAAAAATTTTGGACCAAAATTGCTACCT
ATCATCTATCTTTTTGATTGCAGGGAATCCCTGCCAAGGTAACTTGACAGTCGGCCTAAT

TCTGTTGACAGAAAATGAAGCCTTGACGGTTCTAATTATCCAAAAGTGGGTTTTCATCAG
GACGTACAGTCAGAGTGTGAGTGCATTCTAATGAAAACTTCTTCAGCCCTCATTCAATTG
CATACAAAAGCCCTCAAAGAGAACATACAGTACAGCAGTTTTGTAAAAGGCAACAATACG
ATTTGTACAGACCCCGACACTCCAATCCTATAGATCACCACGTTGCTCCTCTGTCCCCAG
CACCCCTTATTTCTTAAGCAGCACCACCAGTCCCCAGAGGACAAGAGCAGCTATATCTCA
TCCACCTGTGCATTTCTGCACCAGCGATGCANAAAACACCCTGGGGTGGGCCACAGAGAA
AACCGTAACTATGGATACCCAAGGGGC

CA774243

TAAATAACAAACACCACTTTGTTATGAAGACCTTACAAACCTCTTCTTAAGACATTCTTA
CTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTCTTTCCTGACATAAATCCT
TTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGTAATAATTTTTATACTTGG
GAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCTCCGACGGGGCAGGTACTG
TGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTGGTTCCTTCGGTGCCAATG
GTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAGGGCTACACCACTGTCAAC
ATTATCCTGGACTCTGTGTCTCTCTCTGTTGGGTCTTGTGGCATCACATCAGGCCAAAAT
TGCCAGACCAGGACCCTAAGTGTCTGATAGAGGCGATGATCTTTTCCAAAGTCAGTACTT
ACAAACTGGCATTCTTACAGGCTGCACCATTTCCTAGTATGTCTGCTTTAAGCCTGGTTC
AACCTCTCATCGAATATTAAATTTTTCTTTGTA

CA436347

TTTTTTTTTTTTTTTCTTGGGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAAA
AAAATTTCTGTAGGGTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGC
TGAATAAATGAAAATTGGCTCTATTTCTTCAACTTCGGGATAGCCCGAGTAAAAATACTA
ATAATTTCTAAATTTTAGGGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGT
TCAAATACAATACAGTAATTACAAAATATAGACCATCTCTTTACAAATACAAATTATAGT
ATATTACAAGTCATGTACAGTAAATCTATAATTTTAAACAAACTAGTGTATCTAAGTTTA
CCTGGTTGCGAGTGCATTATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGAAA
CCGACCATCGGAGTGATATTCTCTTATGTAAAC

CA389011

TGATTACTTGTAGCAAAGTACTTCCCCACATTTAGCTGGATTTGTCTTTGGTTTGAAGAG
GCTAATACGTGAAAGATTTGTTCACAGTTGGATGTCCCCTTTTCTGAACCATGAAGTAAT
ATTGTGAATGGAGTTGAATGCTGAGGTTAGGGTGCCGGAAAGATTCAGGGTCCTTCGGTA
CCCTCACATGGCTTGGCTTTGGTAGAACAAGAAACTAAGCTCTGATTTGGCTTTAAATGA
GAGTGCTAAATTTCCTTTTTCTAATAAAGAACCTAGCTAAACATTTATATATACTTTTGA
ACACTGAACTNTCTTGTTGCAGAGTTAACAGCTGTTGGGGGTAGCTGACAGCTGGATCCT
GGTGCTGTTGGTACCATGGTACCTGAAGTGCACAGGCTGGTAGCCACACCTGACA

BU679327

TTTTTTTTTTTTTTTCTTACAAAGAAAAATTTAATATTCGATNGAGAGGTTGAACCAGGC
TTAAAGCAGACATACTAGGAAATGGTGCAGCCTGTAAGAATGCCAGTTTGTAAGTACTGA
CTTTGGAAAAGATCATCGCCTCTATCAGACACTTAGGGTCCTGGTCTGGCAATTTTGGCC
TGATGTGATGCCACAAGACCCAACAGAGAGAGACACAGAGTCCAGGATAATGTTGACAGT
GGTGTAGCCCTTTAGGAGAAATGGCGCTCCCTGCGGCTGGTATTAGGTTACCATTGGCAC
CGAAGGAACCAGGAGGATAAGAATATCCATAATTTCAGAGCTGCCCTGGCACAGTACCTG
CCCCGTCGGAGGCTCTCACTGGCAAATGACAGCTCTGTGCAAGGAGCACTCCCAAGTATA
AAAATTATTACACAGTTTTATTCTGAAGAACATTTTGCATTTTAATAAAAAAGGATTTAT
GTCAGGAAAGAGTCATTTACAAACCTTGAAGTGTTTTTGCCTGGATCAGAGTAAGAATGT
CTTAAGAAGAGGTTTGTAAGGTCTTCATAACANAGTGGTGTTTGTTATTTACAAAAAAAA
AAAAAAAAAAAATAAAAAAAAAAAAAAAAACCTCGTGCCGAATTCT

BU608029

TTTTTTTTTTTTTTTTGTAAATAACAAACACCACTTTGGTTATGAAGACCTTACAAACCT
CTTCTTAAGACATTCTTACTCTGATCCAGGCAAAAACACTTCAAGGTTTGTAAATGACTC
TTTCCTGACATAAATCCTTTTTTATTAAAATGCAAAATGTTCTTCAGAATAAAACTGTGT
AATAATTTTTATACTTGGGAGTGCTCCTTGCACAGAGCTGTCATTTGCCAGTGAGAGCCT
CCGACAGGGCAGGTACTGTGCCAGGGCAGCTCTGAAATTATGGATATTCTTATCCTCCTG
GTTCCTTCGGTGCCAATGGTAACCTAATACCAGCCGCAGGGAGCGCCATTTCTCCTAAAG
GGCTACACCACTGTCAACATTATCCTGGACTCTGTGTCTCTCTCTGTTGGGTCTTGTGGC
ATCACATCAGGCCAAAATTGCCAGACCAGGACCCTAAGTGTCTGATAGAGGCGATGATCT
TTTCCAAAGTCAGTACTTACAAACTGGCATTCTTACAGGCTGCACCATTTCCTAGTATGT
CTGCTTTAAGCCTGGTTCAACCTCTCATCGAATATTAAATTTTTCTTTGTAAGAAAAATT
TGAAGTTGTAGAGCATGGTTTTTTGTTTTCCCTTGTCTTAGGAAAGTTTTAAGATGAAAT
GTTTTTCC

BU073743

AGTACACAAGGTGAAACTGCTCCAGTTTTTCTCATAGCAGGGTCAGCAGGAAAGCAAGTG
GTGCCCCTGGTCCCATCTCACACAGGTGAGACTGCACCGAGAGGTAACGTGGCCCTCACA
GCCCACCACGCCTGGCCTTCGCCCAATTCTGAAACTTCGTAGGATAGAGCTGGAAAGTGC
CACATGGTGAAGCGAGATCCAGCTGTCTGGGTGGATGTCGGAGTCCATAGGCTGAGCAGA
GATGGTTCTTAGTGAGGTTCTCGCTGCCAGTTGACGGTGAAATCATAGCTGCCATTTACA
TTTTGTGAGATTATGAAAAACATAAGACTAAAGAAACTAAATGTGTTATTCCTGTGGACA
CAAAAATGTGTGTTTTTCAGATGGGGAGGGGACCAAAAAGGAAAAACATTTCATCTTAAA
ACTTCCCTAAGACAAAGGAAAACAAAAAACCATGCTCTACAACTTCAAATTTTTCTTACA
AAGAAAAATTTAATAT

BE175413

AGCTGAGGAAACAAAACGAGAGAAGAAGATGATGTGTTCAAAAGAAATGGTGCCCTGCTT
GGAAACCATGTCAATCATGTTAATAGTGATAGGAGAGATTCCCTTCAGCAGACCAATACC
ACCCACCGTCCCCTGCATGTCCAAAGGCCTTCAATTCCACCTGCAAGTGATACTGAGAAA

CCGCTGTTTCCTCCAGCAGGAAATTCGGTGTGTCATAACCATCATAACCATAATTCCATA
GGAAAGCAAGTTCCCACCTCAACAAATGCCAATCTCAATAATGCCAATATGTCCAAAGCT
GCCCATGGAAAGCGGCCCAGCATAGGGAACCTTGAGCATGTGTCTGAAAATGGGCATCAT
TCTTCCCACAAGCATGACCGGGAGCCTCAGAGAAGGTCCAGTGTGAAAAGGTCCGACTCA
GGAGATGAACAGCTCCCAACTATTGGCCGGGAAGACCCAGAGATACATGGCTATTTCAGG
CACCCCCACGGCTTGGGGGAGCAGGAGTATTTCAGTAGTGAGGAATGCTACGAGGATGAC
AGCTCGCCCACCTGGAGCAGGCAAAACTATGGCTACTACAGCAGATACCCAGGCAGAAAC
ATCGACTCTGAGAGGCGCGAGGCTACATCATCCCAAGATTCTGGAGGAGATGACTCGCCG
TTTGTATGATCACGAGATCTCAAGAGAGCTATACTCCCACC

AW969248

TCTTGTGGAAAGATGATAGGTTTATAGTGACTCAAAATATTTTAGAAAAATTTCTGTAGG
GTCAAGTTCTTTCAAACTTAAAATTTTAACCCCAGAGGATTTTCGCTGAATAAATGAAAA
TTGGCTCTATTTCTTCTACTTCTGGATAGCCCGAGTAAAAATACTAATAATTTCTAGATT
TTAGTGGGGAACTACAATTATTAGGACCCATGGATATTGCTGCAGTTCAAATACAATACA
GTAATTACAAAATATAGACCATCTCTTTACAAATCCAAATTATAGTATATTACAAGTCAT
GTACCGTAAATCTATTTTAAACAAACTAGGGTATCTAAGTTTACCTGGTTGCAAGTGCAT
TATTATTCCAGTTTACAGTTGCCCTTAGCGTGACAGTCAGAAACCGACCATCGGAGTGAT
ATTCTCTTATGTAAACTGGCGTCACATCACAGAAAACCTTATTTATTTGGGGGAAAGGGT
TTAAAAATGGATATGTTGGGCCCAGTGACTGATAC

AI90811

GGAAAAGATCATCGCCTCTATCAGACACTTAGGGTCCTGGTCTGGCAATTTTGGCCTGAT
GTGATGCCACAAGACCCAACAGAGAGAGACACAGAGTCCAGGATAATGTTGACAGTGGTG
TAGCCCTTTAGGAGAAATGGCGCTCCCTGCGGCTGGTATTAGGTTACCATTGGCACCGAA
GGAACCAGGAGGATAAGAATATCCATAATTTCAGAGCTGCCCTGGCACGGTACCTGCCCC
GTCGGAGGCTCTCACTGG

BF754485

GATGCGTGATGGCTGATCTAGAGGTATCCCATGGACTCTCATCGCAGCTCCTGGTACACA
GACGAGCCCGACATCTCCTACCGGACTTTCACACCAGCCAGCCTGACTGTCCCCAGCAGC
TTCCGGAACAAAAACAGCGACAAGCAGAGGAGTGCGGACAGCTTGGTGGAGGCAGTCCTG
ATATCCGAAGGCTTGGGACGCTATGCAAGGGACCCAAAATTTGTGTCAGCAACAAAACAC
GAGATCGCTGATGCCTGTGACCTCACCATCGACGAGATGGAGAGTGCAGCCAGCACCCTG
CTTAATGGGAACGTGCGTCCCCGAGCCAACGGGGATGTGGGCCCCCTCTCACACCGGCAG
GACTATGAGCTACAGGACTTTGGTCCTGGCTACAGCGACGAAGGGCCAGACCCTGGGAGG
GATGAGGAGGACCTGGCGGATGAAATGATATGCATCACCACCTTGTAGCCCCCAGCGAGG
GGCAGACTGGCTCTGGCCTCAGGTGGGGCG

BI015409

CGCTCGTTCGCTGTGCCAGGACAAAGTCCTGTAGCTCATAGTCCTGCCGTGTGAGAGGGG
GCCACATCCCCGTTNCTCGGGACGCACGACCCATTAAGCAGGGTGCTGGCTGCCCCCTCC
ATCTCGTCGATGGAGAGGTCANCAGGCATCAGCGATTTCGTGTTTTGTGTGCGTGACACA
AATTTTGGGTCCCTTGCATACGCGTCCCACAGCCTTACGGAGTATCAGCGACTGCTCTCC
ACCAATGCTGCCCGCGACTCCTACTGCTTGTCCGCTGTTTTTGGTTCCGGAAGCTGCTGG
GGACAGTCAGGCTGGCTGGTGTGAAAGTCCGGTAGGAGATGTCGGGCTCGTCTGTGTACC
AGGAGCTGCGGTGCAGGGACGGCAGGCTGCCGTTCACCTGGTCCG

BG202552

GAGTTTCGAGCTTCTCTTTTCCTAAGNGAAAAAANAAAGAANCACAAGNAAACCAAATAA
CCATGTTACTCTGTATAAAAATGCTAATCAGGGAATTCTGAATCAATAATGCTCCAATGA
AGGACAGAATTTAATTAGAAACAACACTAACCACAAGAGCCTAGCACAACCCAAACTCAG
AGCTTCCTGGTAATCTCAATGCGATGGATTCATTACACAGACCATCTTATTAAAATTCTC
ATCTGAGAGCTAATCAGCATTGAATGCATCATTTATTTTATGACACCAAAATTAACTGCA
GTGATTCTTTAAGCATGGGGACACGTGACTCCCACTCTCAGCCCCGAGGGATGACAGCCA
AGAGCCTGGCTTCTGCCCAAGATTCCATCCGTTTTGGTCTGCAGTGCATGGTCAACCATG
ATCCACAAAGCAGCAACCCGGGGGCTGTAGCTGCCGTGATGCGGGGGTAAGCCTGGCAGG
CTGCAACTGTTGCAGGGCTCCCAACACAGCCCCTGGACAAACGCGTCAGGGGAAAATAGG
GTTACCTGGCAATCTTTTTCCTCTCCTTTTCTTCCGCTTCTTCTTTCTGAGCAGTGTTCA
GACTTTCAGCATCAGCCAAAGTGTCTACAGCGATGGCCAAGAAGACATTCAGTAGAATAT
CTAATTACAACTTTTTAAGGGCACAACACACTACTAAATGCAACTACGTGCGGCCAACAA
TGGCAACGCCACACACCTCTGCATCCCGGGAAGCTGGGTAGTAGGTGACGTCCCCAAGTG
TTATACTCACACAGCAAACCTAGAGTACCAGAGCCCTGCTTTTCAAACAANACANAACAA
ACAAACAACCCAAAGTAAAACCTGGTAAGGGACGTCTTCAGAAGTAAATTAC

BF883669

CTGGCTTTCCCATAGCACGCTCGGCAGGAAAGCAAGTGATGCCCCTGGCTCCCATCTCAC
ACAGGTGACACTGCACCGAGAGGTAACGTGGCCCTCACAGCCCACCACGCCTGGCCTTCG
CCCAATTCTGAAACTTCGTAGGATAGAGCTGGAAAGTGGCACATGGTGAAGCGAGATCCA
GCTGTCTGGGTGGATGTCGGAGCTCCATAGGCTGAGCAGAGATGGTTCTTAGTGAGGTTC
TCGCTGCCAGTTGACGGTGAAATCATAGCTGCCATTTACATTTTGTGAGATTATGAAAAA
CATAAGACTAAAGAAACTAAATGTGTTATTCCTGTGGACACAAAAATGTGTGTTTTTCAG
ATGGGGAGGGGACCAAAAAGGAAAAACATTTCATCTTAAAACTTTCCTAAGACAAAGGAA
AACAA

BF817590

CTCAGCATGNATGAAACAGGATGAGGTTGGTGAAGATGTGGTGGTTGATGAGCTTGTGGC
AGCCTACGCGGATCGGGTTGGTCTTGCTAAGAATGAAGAAAGCGCTCCCTTCAGGGATGG
GGGCAATTTTTTCCTTCATGTTCAACTCCGAGATCCTTCGAGGACGGGGTCCGGCAGGAA
CCTCAGGTTCATCCTCCTCCTCTTCCTCTTCCTCTTCCCCTACGGGCACATCGCAAGGCG
GATAGGGGTCCTTGTCTTCATCCTCTTCTCTATAGTCATCAATTGTAACCTTGTTGTCAC

TGTTGGCTATCTGGTTGACTTCTGGTTTGTTGTTCTTTTTATTTTCTAGGCTCTCTTTTC
TGGCAATCTTTTTCCTCTCCTTTTCTTCCGCTTCTTCTTTCTGAGCAGTGTTCAGACTTT
CAGCATCAGCCAAATGGTCTA

BF807128

TCAAAGTCGAAGGAGGATCTCCGCGTGGGATGCTGGGGTGGGAGGTAGTAGGCGTCTCCT
TGGAGATCTCCGTGAATCATAGCAAACGGGCGAGTCATCGTCCTACAAGAATCCTAGTGG
ATGATGGTAGCCTCGGGGCCTCTCAGAGTCGATGTTTCTGCCTGG

BF806160

CTCGCCCGTTTGCTATGAGTCACGGAGATCTCCAAGGAGACGCCTACTACCTCCCACCCC
AGCATCCCACCGGAGATCCTCCTTCAACTTTGAGTGCCTGCGCCGGCAGAGCAGCCAGGA
AGAGGTCCCGTCGTCTCCCATCTTCCCCCATCGCACGGCCCTGCCTCTGCATCTAATGCA
GCAACAGATCATGGCAGTTGCCGGCCTAGATTCAAGTAAAGCCCAGAAGTACTCACCGAG
TCACTCGACCCGGCCGTGGGCCACCCCTCCAGCAACCCCTCCCTACCGGGACTGGACACC
GTGCTACACCCCCCAGATGACGCCGATGTA

BF805244

CCAGGCAGAAACATCGACTCTGAGAGGCCCCGAGGCTACCATCATCCCCAAGGATTCTTG
GAGGACGATGACTCGCCCGTTTGCTATGATTCACGGAGATCTCCAAGGAGACGCCTACTA
CCTCCCACCCCAGCATCCCACCGGAGATCCTCCTTCAACTTTGAGTGCCTGCGCCGGCAG
AGCAGCCAGGAAGAGGTCCCGTCGTCTCCCATCTTCCCCCATCGCACGGCCCTGCCTCTG
CATCTAATGCAGCAACAGATCATGGCAGTTGCCGGCCTAGATTCAAGTAAAGCCCAGAAG
TACTCACCGAGTCACTCGACCCGGTCGTGGGCCACCCCTCCAGCAACCCCTCCCTACCGG
GACTGGACACCGTGCTACACCCCCCAGATGACGCCGATGTA

BF805235

TACATCGGCGTCATCTGGGGGGTGTAGCACGGTGTCCAGTCCCGGTAGGGAGGGGTTGCT
GGAGGGGTGGCCCACGACCGGGTCGAGTGACTCGGTGAGTACTTCTGGGCTTTACTTGAA
TCTAGGCCGGCAACTGCCATGATCTGTTGCTGCATTAGATGCAGAGGCAGGGCCGTGCGA
TGGGGGAAGATGGGAGACGACGGGACCTCTTCCTGGCTGCTCTGCCGGCGCAGGCACTCA
AAGTTGAAGGAGGATCTCCGGTGGGATGCTGGGGTGGGAGGTAGTAGGCGTCTCCTTGGA
GATCTCCGTGAATCATAGCANACGGGCGAGTCATCGTCCTCCAAGAATCCTTGNNGATGA
TGGTAGCCTCGGNGCCTCTCAGAGTCGATGTTTCTGCCTGNGTATCTGCTCGGGCGAGCC
GGTACCGAGCT

BF805080

TACATCGGCGTCATCTGGGGGGTGTAGCACGGTGTCCAGTCCCGGTAGGGAGGGGTTGCT
GGAGGGGTGGCCCACGACCGGGTCGAGTGACTCGGTGAGTACTTCTGGGCTTTACTTGAA

TCTAGGCCGGCAACTGCCATGATCTGTTGCTGCATTAGATGCAGAGGCAGGGCCGTGCGA
TGGGGGAAGATGGGAGACGACGGGACCTCTTCCTGGCTGCTCTGCCGGCGCAGGCACTCA
AAGTTGAAGGAGGATCTCCGGTGGGATGCTGGGGTGGGAGGTAGTAGGCGTCTCCTTGGA
GATCTCCGTGAATCATAGCAAACGGGCGAG

T27949

GCGGACAGCTTGGTGGAGGCAGTCCTGATATCCGAAGCCTTNGGACGCTATGCAAGGGAC
CCAAAATTTNTTTCAGCAACAAAACACGAAATCGCTGATGCCTGTAACCTCACCATCGAC
GAGATGGAGAGTNCAGCCAGCACCCTGCTTAATGGGAACGTGCGTCCCCGAGCCAACGGG
GAT

BE836638

AAGAAATAGGAGGATAAGAATATCATATTTCAGAGCTGCCCTGGCACAGTACCTGCCCCG
TCGGAGGCTCTCACTGGCAAATGACAGCTCTGTGCAAGGAGCACTCCCAAGTATAAAAAT
TATTACACAGTT

BE770685

CCATTGGTACGAGAGAAATTAGGAGGATAAGATTATCTATTATTCTGAGCTGCCCTGGCA
CAGTACCTGCCCCGTCGGAGGCTCTCACTGGCAAATGACAGCTCTGTGCAAGGAGCACTC
CCAAGTATAAAAATTATTACATAGTTTTATTCTGAAGAACATTTTGCATTTTAATAAAAA
AGGATTTATGTCAGGAAAGAGTCATTTACATACCTTGAATTGTTTTTGCCTGGATCAGAG
TAAGAATGTCTTAAGAAGAGGTTTGTAAGGTCTTCATAACAAAGTGGTGTTTGTTATTTA
CAAAAAAAAAAAAAAAAAAAATTTTTATACCGGGTTTGTCTGTATACAAATTTCTCTG

BE769065

TCCAGAGTAGAAGAAATCAGCCAAGTATCATTTATTCAGCGAAAATCCTCTGGGGATTAA
AATTTTAAGTTTGAAAGAACTTGACACTACAGAAATTTTTCTAAAATATTTTGAGTCACT
ATAAACCTATCATCTTTCCACAAGATATACCAGATGACTATTTGCAGTCTTTTCTTTGGG
CAAGAGTTCCATGATTTTGATACTGTACCTTTGGATCCACCATGGGTTGCAACTGTCTTT
GGTTTTGTTTGTTTGACTTGAACCACCCTCTGGAAAGCTACTCTGGAAA

Sequences identified as those of HOXB13 cluster

BF676461

GGGATTCCCCCGGCTGGGTGGGGAGAGCGAGCTGGGTGCCCCCATAGATTCCCCTGCCCG
AACCTCATGAGCCGACCCTCGGCTCCATGGAGCCCGGAAATTATGCCACCTTGGATGGAG
CCAAGGATATCGAAGGCTTGTTGGGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCC
TCTCTGACCAGCCACCCAGCGCGCTACGCTTGATGCCTGTGTCAATATGCCCCCTTGATC

TGCCAGGCTCGGGGAGCGGCCAAAAGCAATGCCCACCCTATGCTCTGGGGGTGCCCAGGG
GACTGTCCCCGGCTCCGTGCCTTATGGTTACTGTGGGGCGGGGTACATACTCCTGCAGAG
TTGTCCCGGAGCTCGTTGAAACCTTGTGCCGAGGAGAGCCACCCTGGCGGTACCCGGGAA
GACTCCCCAGGGCGGGAAGAGTACCCCAGCGGCCCAATGAGTTGTGCTTCTATCGGGATA
TCCGGGACCTACCAGGCCTATGTGCAGGTACTGGACGTGTCCTGTGCTGCAGACTCTGGG
TGTCCGTGGAGCACCGGACATTGGCTCGCTGTGGCCTGTGGCCGGTACCAGTCTTGGGCT
CTCGGTGTGTGGCTGGACACGCCGGTTGTGTTCGCGGGAGACCGCACCCACCAGGTTCCT
TTGGGAGGGCCGCTTTGCAGACTCCGGGGGAGGCCCCTCTGAGGCGGGGCCTTTTCGGGG
GGGCGAAGAAAGCTTTCCGACGCAGGCGCTTGCGGAGCTGGCGGGACATCGGGACACTTC
ACCCAGCGAAGCGCGGCTTGGGGCCCCTCTGGGCGCGGTCTCGGTTGACACCGGCGAAGA
GTTTCGGGAGAGGCCCATATCTTCTGGGGAGGGCGTTGCGTCGCCCCCG

BC007092

ggattccccc ggcctgggtg gggagagcga gctgggtgcc ccctagattc cccgcccccg
cacctcatga gccgaccctc ggctccatgg agcccggcaa ttatgccacc ttggatggag
ccaaggatat cgaaggcttg ctgggagcgg gagggggggcg gaatctggtc gcccactccc
ctctgaccag ccacccagcg gcgcctacgc tgatgcctgc tgtcaactat gccccccttgg
atctgccagg ctcggcggag ccgccaaagc aatgccaccc atgccctggg gtgccccagg
ggacgtcccc agctcccgtg ccttatggtt actttggagg cgggtactac tcctgccgag
tgtcccggag ctcgctgaaa ccctgtgccc aggcagccac cctggccgcg taccccgcgg
agactcccac ggccggggaa gagtacccca gccgccccac tgagtttgcc ttctatccgg
gatatccggg aacctaccag cctatggcca gttacctgga cgtgtctgtg gtgcagactc
tgggtgctcc tggagaaccg cgacatgact ccctgttgcc tgtggacagt taccagtctt
gggctctcgc tggtggctgg aacagccaga tgtgttgcca gggagaacag aacccaccag
gtccctttg gaaggcagca tttgcagact ccagcgggca gcaccctcct gacgcctgcg
cctttcgtcg cggccgcaag aaacgcattc cgtacagcaa ggggcagttg cgggagctgg
agcgggagta tgcggctaac aagttcatca ccaaggacaa gaggcgcaag atctcggcag
ccaccagcct ctcggagcgc cagattacca tctggtttca gaaccgccgg gtcaaagaga
agaaggttct cgccaaggtg aagaacagcg ctacccctta agagatctcc ttgcctgggt
gggaggagcg aaagtggggg tgtcctgggg agaccaggaa cctgccaagc ccaggctggg
gccaaggact ctgctgagag gcccctagag acaacaccct tcccaggcca ctggctgctg
gactgttcct caggagcggc ctgggtaccc agtatgtgca gggagacgga accccatgtg
acagcccact ccaccagggt tcccaaagaa cctggcccag tcataatcat tcatcctgac
agtggcaata atcacgataa ccagtactag ctgccatgat cgttagcctc atattttcta
tctagagctc tgtagagcac tttagaaacc gctttcatga attgagctaa ttatgaataa
atttggaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa

BM462617

ATTCCCCCGGCCTGGGTGGGGAGAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCGCA
CCTCATGAGCCGACCCTCGGCTCCATGGAGCCCGGCAATTATGCCACCTTGGATGGAGCC
AAGGATATCGAAGGCTTGCTGGGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCCCT
CTGACCAGCCACCCAGCGGCGCCTACGCTGATGCCTGCTGTCAACTATGCCCCCTTGGAT

CTGCCAGGCTCGGCGGAGCCGCCAAAGCAATGCCACCCATGCCCTGGGGTGCCCCAGGGG
ACGTCCCCAGCTCCCGTGCCTTATGGTTACTTTGGAGGCGGGTACTACTCCTGCCGAGTG
TCCCGGAGCTCGCTGAAACCCTGTGCCCAGGCAGCCACCCTGGCCGCGTACCCCGCGGAG
ACTCCCACGGCCGGGGAAGAGTACCCCAGCCGCCCCACTGAGTTTGCCTTCTATCCGGGA
TATCCGGGAACCTACCAGCCTATGGCCAGTTACCTGGACGTGTCTGTGGTGCAGACTCTG
GGTGCTCCTGGAGAACCGCGACATGACTCCCTGTTGCCTGTGGACAGTTACCAGTCCTGG
GCTCTCGCTGGTGGCTGGAACAGCCAGATGTGTTGCCAGGGAGAACAGAACCCACCAGGT
CCCCTTTTGGAAGGCAGCATTTGCAGACTCCAGCGGGCAGCACCCTCCTGACGCCTGCGC
CTTTCGT

BG752489

GCAGGCGACTTGCGAGCTGGGAGCGATTTAAAACGCTTTGGATTCCCCCGGCCTGGGTGG
GGAGAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCGCACCTCATGAGCCGACCCTCG
GCTCCATGGAGCCCGGCAATTATGCCACCTTGGATGGAGCCAAGGATATCGAAGGCTTGC
TGGGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCCCTCTGACCAGCCACCCAGCGG
CGCCTACGCTGATGCCTGCTGTCAACTATCCCCCCTTGGATCTGCCAGGCTCGGCGGAGC
CGCCAAAGCAATGCCACCCATGCCCTGGGGTGCCCCAGGGGACGTCCCCAGCTCCCGTGC
CTTATGGTTACTTTGGAGGCGGGTACTACTCCTGCCGAGTGTCCCGGAGCTCGCTGAAAC
CCTGTGCCCAGGCAGCCACCCTGGCCGCGTACCCCGCGGAGACTCCCACGGCCGGGGAAG
AGTACCCCAGCCGCCCCACTGAGTTTGCCTTCTATCCGGGATATCCGGGAACCTACCAGC
CTATGGCCAGTTACCTGGACGTGTCTGTGGTGCAGACTCTGGGTGCTCCTGGAGAACCGC
GACATGACTCCCTGTTGCCTGTGGACAGTTACCAGTCTTGGGCTCTCGCTGGTGGCTGGA
ACAGCCAGATGTGTTGCCAGGGAGAACAGAAGCCACCAGGTCCCTTTTGGAAGGCAGCAT
CTGCAGACTCCAGCGGGCAGGACCTCCTGACGCCTGCGGCCTTTCGTCGCGAGCGCAAGA
AACGCATTCCGTA

BG778198

GGATTTAAAACGCTTTGGATTCCCCCGGCCTGGGTGGGGAGAGCGAGCTGGGTGCCCCCT
AGATTCCCCGCCCCCGCACCTCATGAGCCGACCCTCGGCTCCATGGAGCCCGGCAATTAT
GCCACCTTGGATGGAGCCAAGGATATCGAAGGCTTGCTGGGAGCGGGAGGGGGGCGGAAT
CTGGTCGCCCACTCCCCTCTGACCAGCCACCCAGCGGCGCCTACGCTGATGCCTGCTGTC
AACTATGCCCCCTTGGATCTGCCAGGCTCGGCGGAGCCGCCAAAGCAATGCCACCCATGC
CCTGGGGTGCCCCAGGGACGTCCCCAGCTCCCGTGCCTTATGGTTACTTTGGAGGCGGGT
ACTACTCCTGCCGAGTGTCCCGGAGCTCGCTGAAACCCTGTGCCCAGGCAGCCACCCTGG
CCGCGTACCCCGCGGAGACTCCCACGGCCGGGGAAGAGTACCCCAGCCGCCCCACTGAGT
TTGCCTTCTATCCGGGATATCCGGGAACCTACCAGCCTATGGCCAGTTACCTGGACGTGT
CTGTGGTGCAGACTCTGGGTGCTCCTGGAGAACCGCGACATGACTCCCTGTTGCCTGTGG
ACAGTTACCAGTCTTGGGCTCTCGCTGGTGGGCTGGAACAGCCAGATGTGTTGCCAGCGC
AGAACAGAACCCACCAGGTCCCTTTTGGAAGGCAGCATTTGCAGACTCCAGCGGGCAGAA
CCCTCCTGACGCCTGCGCCTTTCGTTCGCGGGCGAAAAA

CB050884

AAGAAACGCATTCCGTACAGCAAGGGGCAGTTGCGGGAGCTGGAGCGGGAGTATGCGGCT
AACAAGTTCATCACCAAGGACAAGAGGCGCAAGATCTCGGCAGCCACCAGCCTCTCGGAG
CGCCAGATTACCATCTGGTTTCAGAACCGCCGGGTCAAAGAGAAGAAGGTTCTCGCCAAG
GTGAAGAACAGCGCTACCCCTTAAGAGATCTCCTTGCCTGGGTGGGAGGAGCGAAAGTGG
GGGTGTCCTGGGGAGACCAGGAACCTGCCAAGCCCAGGCTGGGGCCAAGGACTCTGCTGA
GAGGCCCCTAGAGACAACACCCTTCCCAGGCCACTGGCTGCTGGACTGTTCCTCAGGAGC
GGCCTGGGTACCCAGTATGTGCAGGGAGACGGAACCCCATGTGACAGCCCACTCCACCAG
GGTTCCCAAAGAACCTGGCCCAGTCATAATCATTCATCCTGACAGTGGCAATAATCACGA
TAACCAGTACTAGCTGCCATGATCGTTAGCCTCATATTTTCTATCTAGAGCTCTGTAGAG
CACTTTAGAAACCGCTTTCATGAATTGAGCTAATTATGAATAAATTTGGAAGGCGAAAAA
AAAAACCTCGTGCC

CB050885

ATTCGGCACGAGGTTTTTTTTTTCGCCTTCCAAATTTATTCATAATTAGCTCAATTCATG
AAAGCGGTTTCTAAAGTGCTCTACAGAGCTCTAGATAGAAAATATGAGGCTAACGATCAT
GGCAGCTAGTACTGGTTATCGTGATTATTGCCACTGTCAGGATGAATGATTATGACTGGG
CCAGGTTCTTTGGGAACCCTGGTGGAGTGGGCTGTCACATGGGGTTCCGTCTCCCTGCAC
ATACTGGGTACCCAGGCCGCTCCTGAGGAACAGTCCAGCAACCAGTGGCCTGGGAAGGGT
GTTGTCTCTAGGGGCCTC

BF965191

GGGTGGGGAGAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCGCACCTCATGAGCCGA
CCCTCGGCTCCATGGAGCCCGGCAATTATGCCACCTTGGATGGAGCCAAGGATATCGAAG
GCTTGCTGGGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCCCTCTGACCAGCCACC
CAGCGGCGCCTACGCTGATGCCTGCTGTCAACTATGCCCCCTTGGATCTGCCAGGCTCGG
CGGAGCCGCCAAAGCAATGCCACCCATGCCCTGGGGTGCCCCAGGGGACGTCCCCAGCTC
CCGTGCCTTATGGTTACTTTGGAGGCGGGTACTACTCCTGCCGAGTGTCCCGGAGCTCGC
TGAAACCCTGTGCCAGGCAGCCACCCTGGCCGCGTAACCCGACGGAGACTCTCACGTGCG
GGGAAGAGTACCCCTAGCGCCCCACATGAGTTTGCCTTCTATCCGGGATATCCGGGACCG
TACCAGCCTATGGCAGTTACCTGGACGTGTCTGTGGTGCCGACTCTGGGTGCTCCTGGAG
AACCGCGGACATGACTCCTTGTTTGCTGTGCGACGCTCACCAGTCTGGGCTCCTCGTCGG
TGGTCGCACTCCCACTTTTTGCCGGGCGACATCCCCCGGGGCCCCTTCCGGAACAGCGAC
CTTGCGAGCCCCCGGGGACACACCCCCGTAAGCGGCCTATCATCGCTGATAAACCTCATC
AGAGGGCACCGAAAGCCGCGACTCTAACCCCCCCACTACGACTCACGACCGCACAGGTAC
TCGAACCGCCCAATATCTGGTTCTAACCCATGGCGCATCTCAGCCGCTAGAGAGCCAACC
AAACGCGCCACGCGCAACCACACTACACCACGGCACCCCTTTCATCTCACTCCCACGCCG
ATCACTCTTCACCCTCCAGAATCATTCCCCTCGCACATCCTACCTATCTCATGCCTCCCA
GTTCACCCCATTCCCTCCCCTAATCTCACCCACACATTCACGCACGTTCTCACTACGCTT
CGCTCCGACCCACATCCTCACCCCCACATTCATACCACTTCACCATCACGACCCCCCCCT
CTCATCGACTCCTGTCTCATTCTCAACCACAGTACTACCAGCTCCAACACACCACTCACC
CCAAGCTATCCATCACCTACACGCTTTCACCCCTCACCGCTCCCAAGTAATTCAGATCAC
TCAAACACAATCTGCTACATACTCATCCCTCCCCCACTCCCAGTACAGTCCAACCACCGA

CCAACTACCTCCGCGCCACCCGCGCCGCCCCACCTCACCGGCCCCAACCGCCCGCACAGG
GCACGCACCCCCCGGCAACCGCGCGATCCGGCCGTACACACTCTTGGGCGGCACGCAGCT
GAGGACATTCCGCGGGAGCGCCCCACCGTGGGCTACGTGGGTCGCGACCCGGCGGGGCGC
GTGCGGCGTCGCCCGCCCGCCCGCCGACTGCGACCCAGTCGAG

BU930208

GGGGCTTTGGATTCCCCCGGCCTGGGTGGGGAGAGCGAGCTGGGTGCCCCCTAGATTCCC
CGCCCCCGCACCTCATGAGCCGACCCTCGGCTCCATGGAGCCCGGCAATTATGCCACCTT
GGATGGAGCCAAGGATATCGAAGGCTTGCTGGGAGCGGGAGGGGGGCGGAATCTGGTCGC
CCACTCCCCTCTGACCAGCCACCCAGCGGCGCCTACGCTGATGCCTGCTGTCAACTATGC
CCCCTTGGATCTGCCAGGCTCGGCGGAGCCGCCAAAGCAATGCCACCCATGCCCTGGGGT
GCCCCAGGGGACGTCCCCAGCTCCCGTGCCTTATGGTTACTTTGGAGGCGGGTACTACTC
CTGCCGAGTGTCCCGGAGCTCGCTGAAACCCTGTGCCCAGGCAGCCACCCTGGCCGCGTA
CCCCGCGGAGACTCCCACGGCCGGGGAAGAGTACCCCAGCCGCCCCACTGAGTTTGCCTT
CTATCCGGGATATCCGGGAACCTACCAGCCTATGGCCAGTTACCTGGACGTGTCTGTGGT
GCAGACTCTGGGTGCTCCTGNAGAACCGCGACATGACTCCCTGTTGCCTGTGGACAGTTA
CCAGTCTTGGGCTCTCGCTGGTGGCCTGGAACAGCCCAGATGTGTTTGCCCAGGGNAGAA
CACGAACCCCACCCGGTTCCCCCTTTTGGGAAAGGGCAGCCATTTTGGCCAGCCTTCCAA
GCGGGGCCAACCACCCCCTCCCCTGGACAGGCCCTGGT

AA807966

GCGGCCGCAAGAAACGCATTCCGTACAGCAAGGGGCAGTTGCGGGACTGGAGCGGGAGTA
TGCGGCTAACAAGTTCATCACCAAGGACAAGAGGCGCAAGATCTCGGCAGCCACCAGCCT
CTCGGAGCGCCAGATTACCATCTGGTTTCAGAACCGCCGGGTCAAAGAGAAGAAGGTTCT
CGCCAAGGTGAAGAACAGCGCTACCCCTTAAGAGATCTCCTTGCCTGGGTGGGAGGAGCG
AAAGTGGGGGTGTCCTGGGGAGACCAGGAACCTGCCAAGCCCAGGCTGGGGCCAAGGACT
CTGCTGAGAGGCCCCTAGAGACAACACCCTTCCCAGGCCACTGGCTGCTGGACTGTTCCT
CAGGAGCGGCCTGGGTACCCAGTATGTGCAGGGAGACGGAACCCCATGTGACAGCCCATT
CCACCAGGGTTCCCAAAGAACCTGGCCCAGTCATAATCATTCATCCTGACAGTGGC

AI884491

AGCGGCCGCAAGAAACGCATTCCGTACAGCAAGGGGCAGTTGCGGGAGCTGGAGCGGGAG
TATGCGGCTAACAAGTTCATCACCAAGGACAAGAGGCGCAAGATCTCGGCAGCCACCAGC
CTCTCGGAGCGCCAGATTACCATCTGGTTTCAGAACCGCCGGGTCAAAGAGAAGAAGGTT
CTCGCCAAGGTGAAGAACAGCGCTACCCCTTAAGAGATCTCCTTGCCTGGGTGGGAGGAG
CGAAAGTGGGGGTGTCCTGGGGAGACCAGGAACCTGCCAAGCCCAGGCTGGGGCCAAGGA
CTCTGCTGAGAGGCCCCTAGAGACAACACCCTTCCCAGGCCACTGGCTGCTGGACTGTTC
CTCAGGAGCGGCCTGGGTACCCAGTATGTGCAGGGAGACGGAACCCCATGTGACAGCCCA
CTCCACCAGGGTTCCCAAAGAACCTGGCCCAGTCATAATCATTCATCCTGACAGTGGCAA
TAATCACGATAACCAGTACTAGCTGCCATGATCGTTAGCCTCATATTTTCTATCTAGAGC
TCTGTAGAGCAC

AA652388

GCGGCCGCAAGAAACGCATTCCGTACAGCAAGGGGCAGTTGCGGGACTGGAGCGTGAGTA
TGCGGCTAACAAGTTCATCACCAAGGACAAGAGGCGCAAGATCTCGGCAGCCACCAGCCT
CTCGGAGCGCCAGATTACCATCTGGTTTCAGAACCGCCGGGTCAAAGAGAAGAAGGTTCT
CGCCAAGGTGAAGAACAGCGCTACCCCTTAAGAGATCTCCTTGCCTGGGTGGGAGGAGCG
AAAGTGGGGGTGTCCTGGGGAGACCAGGAACCTGCCAAGCCCAGGCTGGGGCCAAGGACT
CTGCTGAGAGGCCCCTAGAGACAACACCCTTCCCAGGCCACTGGCTGCTGGACTGTTCCT
CAGGAGCGGCCTGGGTACCCAGTATGTGCAGGGAGACGGAACCCCATGTGACAGCCCACT
CCACCAGGGTTCCCAAAGAACCTGGCC

BF446158

TTTTTTTTTTTTTTTTTTTCGCCTTCCAAATTTATTCATAATTAGCTCAATTCATGAAAG
CGGTTTCTAAAGTGCTCTACAAAGCTCTAAATAAAAAATATGAGGCTAACGATCATGGCA
GCTAGTACTGGTTATCGGGATTATTGCCACTGTCAGGATGAATGATTATGACTGGGCCAG
GTTCTTTGGGAACCCTGGTGGAGTGGGCTGTCACATGGGGTTCCGTCTCCCTGCACATAC
TGGGTACCCAGGCCGTTCCTGAGGAACAGTCCACCACCCAGTGGCCTGGGAAGGGTGTTG
TCTCTAGGGGCCTCTCAACAAAGTCCTTGGCCCCAGCCTGGGCTTGGCAGGTTCCTGGTC
TCCCCAGGACACCCCCACTTTCGCTCCTCCCACCCAGGCAAGGAGATCTCTTAAGGGG

AA657924

GACGCNAGGTATGCGGCTAACAAGTTCATCACCAAGGACAAGAGGCGCAAGATCTCGGCA
GCCACCAGCCTCTCGGAGCGCCAGATTACCATCTGGTTTCAGAACCGCCGGGTCAAAGAG
AAGAAGGTTCTCGCCAAGGTGAAGAACAGCGCTACCCCTTAAGAGATCTCCTTGCCTGGG
TGGGAGGAGCGAAAGTGGGGGTGTCCTGGGGAGACCAGGAACCTGCCAAGCCCAGGCTGG
GGCCAAGGACTCTGCTGAGAGGCCCCTAGAGACAACACCCTTCCCAGGCCACTGGCTGCT
GGACTGTTCCTCAGGAGCGGCCTGGGTACCCATGTATGTGCAGGGAGACGGAACCCCATG
TGACAGCCCACTCCACCAGNGTTCCTAAAGAACCCTGGCCAGTCA

AA644637

GCAGGCGACTTGCGAGCTGGGAGCGGTTTAAAACGCTTTGGATTCCCCCGGCCTGGGTGG
GGAGAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCGCACCTCATGAGCCGACCCTCG
GTCCATGGACACGGCAATTATGCCACCTTGGATGGAGCCAAGGATATCGAAGGCTTGCTG
GGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCCCTCTGACCAGCCACCCAGCGGCG
CCTACGCTGATGCCTGCTGTCAACTATGCCCCCTTGGATCTGCCAGGCTCGGCGGACTCT
NAAAGCATATGCCACCCNATGCCCTGGGGTGCCCCAGGGGAACGTCCCCAGCTCCCGTGC
CTTATGGTT

BF222357

GCGGCCGCAAGAAACGCATTCCGTACAGCAAGGGGCAGTTGCGGGAGCTGGAGCGGGAGT
ATGCGGCTAACAAGTTCATCACCAAGGACAAGAGGCGCAAGATCTCGGCAGCCACCAGCC
TCTCGGAGCGCCAGATTACCATCTGGTTTCAGAACCGCCGGGTCAAAGAGAAGAAGGTTC
TCGCCAAGGTGAAGAACAGCGCTACCCCTTAAGAGATCTCCTTGCCTGGGTGGGAGGAGC
GAAAGTGGGGGTGTCCTGGGGAGACCAGGAACCTGCCAAGCCCAGGCTGGGGCCAAGGAC
TCTGCTGAGAGGCCCCTAGAGACAACACCCTTCCCAGGCCACTGGCTGCTGGACTGTTCC
TCAGGAGCGGCCTG

AA527613

GTCGACGAACAGCGCTACCCCTTAAGAGATCTCCTTGCCTGGGTGGGAGGAGCGAAAGTG
GGGGTGTCCTGGGGAGACCGGGAACTGCCAAGCCCAGGCTGGGGCAAGGACTCTGCTGAG
AGGCCCCTAGAGACAACACCCTTCCCAGGCCACTGCTGCTGGACTGTTCCTCAGGAGCGG
CCTGGGTACCCAGTATGTGCAGGGAGACGGAACCCCATGTGACAGCCCACTCCACCAGGG
TTCCCAAAGAACCTGGCCCAGTCATAATCATTCATCCTGACAGTGGCAATAATCACGATA
ACCAGTACTCAGCTGCCATGATCGTTAGCCTCATATT

AA533227

GCGTCGACCCCTTGAAGAGATCTCCTTGCCTGGGTGGGAGGAGCGAAAGTGGGGGTGTCC
TGGGGAGACCAGGAACCTGCCAAGCCCAGGCTGGGGCCAAGGACTCTGCTGAGAGGCCCC
TAGAGACAACACCCTTCCCAGGCCACTGGCTGCTGGACTGTTCCTCAGGAGCGGCCTGGG
TACCCAGTATGTGCAGGGAGACGGAACCCCATGTGACAGCCCACTCCACCAGGGTTCCCA
AAGAACCTGGCCCAGTCATAATCATTCATCCTGACAGTGGCAATAATCACGATAACCAGT
ACTAGCTGCCATGATCGTTAGCCTCATATTTTCTATCTAGAGCTCTGTAGAGCACTTGTA
GAAACCGCTTTCATGAATTGAGCTAATTATGAATAGATTTGGAAGGGGAAAAAAGTGGAA
AAAGTTTTGCCCAAAGTGGGTCGTTTACGTCG

AA456069

CTCCCTGGCAACACATCTGGCTGTTCCAGCACCAGCGAGACCCAAGACTGGTAACTGTCC
ACAGGCAACAGGGAGTCATGTCGCGGTTCTCCAGGAGCACCCAGAGTCTGCACCACAGAC
ACGTCCAGGTAACTGGCCATAGCTGAGTAGGTTCCCGGATATCCCGGATAGAAGGCAAAC
TCAGTGGGGCGGCTGGGGTACTCTTCCCCGGCCGTGGAGAGTCTCCGCGGGGTACGGCCC
AGGGTGGCTGCCTGGGCATCAGGGTTTCAGCGAGCTCCGGGACACTCGGCAGGAGTAGTA
CCCGCCTCCAAAGTAACCATAAGGCACGGGAGCTGGGGACGTCCCTGGGGCACCCCAG

AA455572

TTTAAAACGCTTTGGATTCCCCCGGCCTGGGTGGGGAGAGCGAGCTGGGTGCCCCCTAGA
TTCCCCGCCCCCGCACCTCATGAGCCGACCCTCGGTCCATGGAGCCGGCGAATTATGCCA
CCTTGGATGGAGCCAAGGATATCGAAGGCTTGCTGGGAGCGGGAGGGGGGCGGAATCTGG
TCGCCCACTCCCCTCTGACCAGCCACCCAGCGGCGCTACGTGATGCCTGCTGTCAACTAT
GCCCTTGGATCTGCCAGCTCGCGGAGCCAAAGCAATGCCACCCATGCCCTGGGGTGCCCC

AGGTGACGTCCCCAGCTCCCGTGCCTTATGGTTACTTTGGAGGCGGGTACTACTCCTGCC
GAGTGTCCCGGAGCTCGCTGAAACCCTGTGCCCAGGCAGCCACCCTGGCCGCGTACCCCG
CGATGACTCCCACGGCCGGGGAAGAGTACCCCAGCCGCCCCACTGAGTTTGCCT

BX117624

CAGGCGACTTGCGAGTCTGGGAGCGATTTAAAACGCTTTGGATTCCCCCGGCCTGGGTGG
GGAGAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCGCACCTCATGAGCCGACCCTCG
GCTCCATGGAGCCCGGCAATTATGCCACCTTGGATGGAGCCAAGGATATCGAAGGCTTGC
TGGGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCCCTCTGACCAGCCACCCAGCGG
CGCCTACGCTGATGCCTGCTGTCAACTATGCCCCCTTGGATCTGCCAGGCTCGGCGGAGC
CGCCAAAGCAATGCCACCCATGCCCTGGGGTGCCCCAGGGGACGTCCCCAGCTCCCGTGC
CTTATGGTTACTTTGGAGGCGGGTACTACTCCTGCCGAGTGTCCCGGAGCTCGCTGAAAC
CCTGTGCCCAGGCAGCCACCCTGGCCGCGTACCCCGCGGAGACTCCCACGGCCGGGGAAG
AGTACCCCAGCCGCCCCACTGAGTTTGCCTTCTATCCGGGATATCCGGGAACCTACCAGC
CTATGGCCAGTTACCTTGGACGTGTCTGTGGTGCAGACTCTGGGTGCTCCTGGAGAACCG
CGACATGACTCCCTGNTGCCTGTGGACAGTTACCAGTCTTGGGCTCTCGCTGGTGGCTGG
AACAGCCAGATGTGTTGNCAGGGAGAACAGAACCCACCAGGTCCCTTTTGGAAGGCAGAT
TTGCAGACTNCAGCGGGCA

BQ673782

AGGCAGCCACCCTGGCCGCGTACCCCGCGGAGACTCCCACGGCCGGGGAAGAGTACCCCA
GCCGCCCCACTGAGTTTGCCTTCTATCCGGGATATCCGGGAACCTACCAGCCTATGGCCA
GTTACCTGGACGTGTCTGTGGTGCAGACTCTGGGTGCTCCTGGAGAACCGCGACATGACT
CCCTGTTGCCTGTGGACAGTTACCAGTCTTGGGCTCTCGCTGGTGGCTGGAACAGCCAGA
TGTGTTGCCAGGGAGAACAGAACCCACCAGGTCCCTTTTGGAAGGCAGCATTTGCAGACT
CCAGCGGGCAGCACCCTCCTGACGCCTGCGCCTTTCGTCGCGGCCGCAAGAAACGCATTC
CGTACAGCAAGGGGCAGTTGCGGGAGCTGGAGCGGGAGTATGCGGCTAACAAGTTCATCA
CCAAGGACAAGAGGCGCAAGATCTCGGCAGCCACCAGCCTCTCGGAGCGCCAGATTACCA
TCTGGTTTCAGAACCGCCGGGTCAAAGAGAAGAAGGTTCTCGCCAAGGTGAAGAACAGCG
CTACCCCTTAAGAGATCTCCTTGCCTGGGTGGGAGGAGCGAAAGTGGGGGTGTCCTGGGG
AGACCAGGAACCTGCCAAGCCCCAGGCTGGGGCCAAGGACTCTGCTGAGAGGCCCCTAGA
GACAACACCCTTCCCAGGCCACTGGCTGCTGGACTGTTCCTCAGGAGCGGCCTGAGTACC
CCGTATGTGCAGGGGAGACGGAACCCCCTGTGACCAGCCCCCCTCCACCCGTGGTCTCCC
AGATAACCTGGCCCCCACTCATAAATCATTTCTTCCCGGGCCGGGGGCCAATCATTCCCC
GAACTACCCCGGTACCTTATACAATTAGATTGGACATGAATCCTCTCGGGGGCATTCCCT
ATGGCGCTGAGGCCCCTCACACCT

AI814453

GGGTGCTGTCCTCTGGAGTCTGCAAATGCTGCCTTCCAAAAGGGACCTGGTGGGTTCTGT
TCTCCCTGGCAACACATCTGGCTGTTCCAGCCACCAGCGAGAGCCCAAGACTGGTAACTG
TCCACAGGCAACAGGGAGTCATGTCGCGGTTCTCCAGGAGCACCCAGAGTCTGCACCACA

GACACGTCCAGGTAACTGGCCATAGGCTGGTAGGTTCCCGGATATCCCGGATAGAAGGCA
AACTCAATGGGGCGGCTGGGGTACTCTTCCCCGGCCGTGGGAGTCTCCGCGGGGTACGCG
GCCAGGGTGGCTGCCTGGGCACAGGGTTTCAGCGAGCTCCGGGACACTCGGCAGGAGTAG
TACCCGCCTCCAAAGTAACCATAAGGCACGGGAGCTGGGGACGTCCCCTGGGGCACCCCA
NGGCATGGGTGGCATTGCTTTGGCGGCTCCGCCGAGCCTGGCAGATCCAAGGGGGCATAG
TTGACAGCAGGCATCAGCGTAGGCGCCGCTGGGTGGCTGGTCAAAAGGGAGTGGCGACCA
NATTCCGCCCCCCTCCCGCTTCCCAG

AI417272

GGGTGCTGCCCGCTGGAGTCTGCAAATGCTGCCTTCCAAAAGGGACCTGGTGGGTTCTGT
TCTCCCTGGCAACACATCTGGCTGTTCCAGCCACCAGCGAGAGCCCAGGACTGGTAACTG
TCCACAGGCAACAGGGAGTCATGTCGCGGTTCTCCAGGAGCACCCAGAGTCTGCACCACA
GACACGTCCAGGTAACTGGCCATAGGCTGGTAGGTTCCCGGATATCCCGGATAGAAGGCA
AACTCAGTGGGGCGGCTGGGGTACTCTTCCCCGCCGTGGGAGTCTCCGCGGGGTACGCGG
CCAGGGTGGCTGCCTGGGCACAGGGTTTCAGCGAGCTCCGGGACACTCGGCAGGAGTAGT
ACCCGCCTCCAAAGTAACCATAAGGCACGGGAGCTGGGGACGTCCCCTGGGGCACCCCAG
GGCATGGGTGGCATTGCTTTGGCGGCTCCGCCGAGCCTGGCAGATCCAAGGNGGCATAGT
TGACAGCAGGCATCAGCGTANGCGCCGCTGGGTGGCTGTCAAGAGG

AA535663

TCGACGTTACCTGGACGTGTCTGTGGTGCAGACTCTGGGTGCTCCTGGAGAACCGCGACA
TGACTCCCTGTTGCCTGTGGACAGTTACCAGTCTTGGGCTCTCGCTGGTGGCTGGAACAG
CAGATGTGTTGCCAGGGAGAACAGAACCCACCAGGTCCCTTTTGGAAGGCAGCATTTGCA
GACTCCAGCGGGCAGCACCCTCCTGACGCCTGCGCCTTTCGTCGCGGCCGCAAGAAACGC
ATTCCGTACAGCAAGGGGCAGTTGCGGGACTGGAGCGGGAGTATGCGGCTAACAAGTTCA
TCACCAAGGACAAGAGGCGCAAGATCTCGGCAGCCACCAGCCTCTCGGAGCGCCAGATTA
CCATCTGGTTTCAGAACCGCCGGGTCAAAGAGAAGAAGGTTCTCGCCAAGGTGAAGAACA
GCGCTACCCCTTAAGAGATCTCCTTGCCTGGGTGGGAGGAGCGAAAGTGTG

AI400493

GTCAGGAGGGTGCTGCCCGCTGGAGTCTGCAAATGCTGCCTTCCAAAAGGGACCTGGTGG
GTTCTGTTCTCCCTGGCAACACATCTGGCTGTTCCAGCCACCAGCGAGAGCCCAGGACTG
GTAACTGTCCACAGGCAACAGGGAGTCATGTCGCGGTTCTCCAGGAGCACCCAGAGTCTG
CACCACAGACACGTCCAGGTAACTGGCCATAGGCTGGTAGGTTCCCGGATATCCCGGATA
GAAGGCAAACTCAGTGGGGCGGCTGGGGTACTCTTCCCCGGCCGTGGGAGTCTCCGCGGG
GTACGCGGCCAGGGTGGCTGCCTGGGCACAGGGTTTCAGCGAGCTCCGGGACACTCGGCA
TGAGTAGACCCGCCTTCCAAGTAACCATAAGGCACGGGAGCTGGTAACGTCCCCTGGGGC
ACCCCANGGCCATGGGTGCATTGCTTTGGCGGCTCCGCCGAGCCCTGCAGATCCAAGGTG
GGCATATTGACAGCAGGCATTCACGTATGCGCCCCCTGGGTGGCTGTCATATTGGGGATT
GCGAC

AW779219

GCAGGCGTCAGGAGGGTGCTGCCCGCTGGAGTCTGCAAATGCTGCCTTCCAAAAGGGACC
TGGTGGGTTCTGTTCTCCCTGGCAACACATCTGGCTGTTCCAGCCACCAGCGAGAGCCCA
AGACTGGTAACTGTCCACAGGCAACAGGGAGTCATGTCGCGGTTCTCCAGGAGCACCCAG
AGTCTGCACCACAGACACGTCCAGGTAACTGGCCATAGGCTGGTAGGTTCCCGGATATCC
CGGATAGAAGGCAAACTCAGTGGGGCGACTGGGGTACTCTTCCCGGCCGTGGGGAGTCTC
CGCGGGGTACGCGGCCAGGGGTGGCTGCCTGGGCACCAGGGGTTTCAGCGAGCTCCGGGA
CACTCNGCAGGAAANTAGTACCCGCCTCCCAAAGTAACCATAAGCACCGGACTGNGGGNN
GGACGTCCCCTGGGGCAC

AA594847

GCGACCGGACGAAAGGAGGCGTCAGGAGGGTGCTGCCCGCTGGAGTCTGCAAATGCTGCC
TTCCAAAAGGGACCTGGTGGGTTCTGTTCTCCCTGGCAACACATCTGGCTGTTCCAGCAC
CAGCGAGACCCAAGACTGGTAACTGTCCACAGGCAACAGGGAGTCATGTCGCGGTTCTCC
AGGAGCACCCAGAGTCTGCACCACAGACACGTCCAGGTAACTGGCCATAGCTAGGTAGGT
TCCCGGATATCCCGGATAGAAGGCAAACTCAGTGGGGCGACTGGGGTACTCTTCCCCGGC
CGTGGGAGTCTCCGCGGGGTACGCCCATGGGTGGCTGCCTGGGCACAGGGTTTCAGCGAG
CTCCGGGACA

AI150430

GCAGGCGTCAGGAGGGTGCTGCCCGCTGGAGTCTGCAAATGCTGCCTTCCAAAAGGGACC
TGGTGGGTTCTGTTCTCCCTGGCAACACATCTGGCTGTTCCAGCCACCAGCGAGAGCCCA
AGACTGGTAACTGTCCACAGGCAACAGGGAGTCATGTCGCGGTTCTCCAGGAGCACCCAG
AGTCTGCACCACAGACACGTCCAGGTAACTGGCCATAGGCTGGTAGGTTCCCGGATATCC
CGGATAGAAGGCAAACTCAGTGGGGCGACTGGGGTACTCTTCCCCGGCCGTGGGAGTCTC
CGCGGGGTACGCGGCCAGGGTGGCTGCCTGGGCACAGGGTTTCAGCGAGCTCCGGGACAC
TCGGCAGGAGTAGTACCCGCCTCCAAAGTAACCATAAGGCACGGGAGCTGGATGCGTCCC
CTAGGGCACCCCATGGCATGGGTGGCATTGCTTTGGCGGCTCCGCCGAGCCTGGCAGATC
CAAGGAGGCACTGTT.

AA494387

GGGTGCTGCCCGCTGGAGTCTGCAAATGCTGCCTTCCAAAAGGGACCTGGTGGGTTCTGT
TCTCCCTGGCAACACATCTGGCTGTTCCAGCCACCAGCGAGACCCAAGACTGGTAACTGT
CCACAGGCAACAGGGAGTCATGTCGCGGTTCTCCAGGAGCACCCAGAGTCTGCACCACAG
ACACGTCCAGGTAACTGGCCATAGGCTGGTAGGTTCCCGGATATCCCGGATAGAAGGCAA
ACTCAGTGGGGCGGCTGGGGTACTCTTCCCCGGCCGTGGGAGTCTCCGCGGGGTACGCGT
CCAGGGTGGCTGCCTGGGCACAGGGTTTCAGCGAGCTCCGGGACACTCGGCAGGAGTAGT
ACCCGCCTCCAAAGTAACCATAAGGCACGGGAGCTGGGGACGTCCCTG

AA662643

GGGTGCTGCCCGCTGGAGTCTGCAAATGCTGCCTTCCAAAAGGGACCTGGTGGGTTCTGT
TCTCCCTGGCAACACATCTGGCTGTTCCAGCCACCAGCGAGACCCAAGACTGGTAACTGT
CCACAGGCAACAGGGAGTCATGTCGCGGTTCTCCAGGAGCACCCAGAGTCTGCACCACAG
ACACGTCCAGGTAACTGGCCATAGGTGGTAGGTTCCCGGATATCCCGGATAGAAGGCAAA
CTCAGTGGGGCGGCTGGGGTACTCTTCCCCGGCCGTGGGAGTCTCCGCGGGGTACGCGGC
CAGGGTGGCTGCCTGGGCACAGGGTTTCAGCGAGCTCCGGGACA

AI935940

GGGTGCTGCCCGCTGGAGTCTGCAAATGCTGCCTTCCAAAAGGGACCTGGTGGGTTCTGT
TCTCCCTGGCAACACATCTGGCTGTTCCTGCCACCAGCGAGAGCCCAAGACTGGTAACTG
TCCACAGGCAACAGGGAGTCATGTCGCGGTTCTCCAGGAGCACCCAGAGTCTGCACCACA
GACACGTCCAGGTAACTGGCCATAGGCTGGTAGGTTCCCGGATATCCCGGATAGAAGGCA
AACTCAGTGGGGCGGCTGGGGTACTCTTCCCCGGCCGTGGGAGTCTCCGCGGGGTACGCG
GCCAGGGTGGCTGCCTGGGCACAGGGTTTCAGCG

AA532530

GGGTGCTGCCCGCTGGAGTCTGCAAATGCTGCCTTCCAAAAGGGACCTGGTGGGTTCTGT
TCTCCCTGGCAACACATCTGGCTGTTCCAGCCACCAGCGAGACCCAAGACTGGTAACTGT
CCACAGGCAACAGGGAGTCATGTCGCGGTTCTCCAGGAGCACCCAGAGTCTGCACCACAG
ACACGTCCAGGTAACTGGCCATAGGTNGGTAGGTTCCCGGATATCCCGGATAGAAGGCAA
ACTCAGTGGGGCGGCTGGGGTACTCTTCCCCGGCCGTGGGAGTCTCCG

AA857572

CTCCCTGGCAACACATCTGGCTGTTCCAGCACCAGCGAGAGCCAAGACTGGTAACTGTCC
ACAGGCAACAGGGAGTCATGTCGCGGTTCTCCAGGAGCACCCAGAGTCTGCACCACAGAC
ACGTCCAGGTAACTGGCCATAGGTCGGTAGGTTCCCGGATATCCCGGATAGAAGGCAAAC
TCAGTGGGGCGACTGGGGTACTCTTCCCCGGCCGTGGGAGTCTCCGCGGGGTACGGCNAC
AGGGTGGCTGCCTGGGCACAGGGTTTCAGCGAGCTCCGGGACACTCGGCAGGAGTAGTAN
CCGCCTCAAAGTAACCATAANGCACGGGAGCTGGGGACGTCCC

AI261980

ACGAAAGGCGCAGGCGTCAGGAGGGTGCTGCCCGCTGGAGTCTGCAAATGCTGCCTTCCA
AAAGGGACCTGGTGGGTTCTGTTCTCCCTGGCAACACATCTGGCTGTTCCAGCCACCAGC
GAGAGCCCAAGACTGGTAACTGTCCACAGGCAACAGGGAGTCATGTCGCGGTTCTCCAGG
AGCACCCAGAGTCTGCACCACAGACACGTCCAGGTAACTGGCCATAGGCTGGTAGGTTCC
CGGATATCCCGGATAGAAGGCAAACTCAGTGGGGCGACTGGGGTACTCTTCCCCGGCCCG
GGGAGTCTCCGCGGGGTACGCGGCCAGGGTGGCTGCCTGGGCACAGGGTTTCAGCGAGCT
CCGGGACACTCGGCGGAGNTAGTACCCGCCTCCAAAGTAACCATAAGGCACGGGAGCTGG
GGAACCGTCCCCTGGGGCACC

BE888751.1

GAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCGCACCTCATGAGCCGACCCTCGGCT
CCATGGAGCCCGGCAATTATGCCACCTTGGATGGAGCCAAGGATATCGAAGGCTTGCTGG
GAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCCCTCTGACCAGCCACCCAGCGGCGC
CTACGCTGATGCCTGCTGTCAACTATGCCCCCTTGGATCTGCCAGGCTCGGCGGAGCCGC
CAAAGCAATGCCACCCATGCCCTGGGGTGCCCCAGGGACGTCCCCAGCTCCCGTGCCTTA
TGGTTACTTTGGAGGCGGGTACTACTCCTGCCGAGTGTCCCGGAGCTCGCTGAAACCCTG
TGCCCAGGCAGCCACCCTGGCCGCGTACCCCGCGGAGACTCCCACGGCCGGGGAAGAGTA
CCCCAGCCGCCCCACTGAGTTTGCCTTCTATCCGGGATATCCGGGAACCTACCAGCCTAT
GGCCAGTTACCTGGACGTGTCTGTGGTGCAGACTCTGGGTGCTCCTGGAGAACCGCGACA
TGACTCCCTGTTGCCTGTGGACAGTTACCAGTCTTGGGCTCTCGCTGGTGGCTGGAACAG
CCAGATGTGTTGCCAGGGAGAACAGAACCCACCAGGTCCCTTTTTGGAAGGCAGCATTTG
CAGACTCCAGCGGCAGGACCTCCTGAACGCCTGCGCCTTTCGTCGCGGCGTCTAAAGTAA
TCCTCGAGG

AI378797

GCGGCCGCGGCCCACCACCAACTGCTCGCCACCGACCCCACTACTCGCCACCGACCCGCT
GCTCGGAGCTTCGGTTCTGCGGGTTGTCCAGACTTCAGGCCTGTGCGCTCAATCGTGGAG
AATGCGCCGGCAGGCCCCCCACCCCCAGCCTAAGGTGCAGGAAGGACCAGCACGAACCCG
CTGGCTTTGCTGCGCGGCCAGGAGATGAGTCCCACCGGGCACTGAGCCCAGGTACAGGAC
ATCAGAGAATGAACACAGAGGCAGAGGCCCTCATGTCCCTCTCAGAGTCCCGGCTCTGCA
NAGAGCCCGTCTGTCTCCAGCTTCCAGAATTCCGCACTGTGAATCTGTCTACGTGGACTG
GGAAAACAGGGTTGGCACCACTCTGCCACTCCGTTTGTGCCTGGGAAGGGCTAAGTATGC
AAGGCTACAAACATCTACTTCACTGGGATCCCAAATGCTCAACAAACCATGACCTGCTNT
GGTCAGAACCACCAGAAATATT

AA234220

GCAGGCGACTTGCGAGCTGGGAGCACTTTAAAACGCTTTGGATTCCCCCGGCCTGGGTGG
GGAGAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCGCACCTCATGAGCCGACCCTCG
GCTCCATGGAGCCTGGCATATTATGCCACCTTGGTATGGAGCCAAGGATATCGAAGGCTT
GCTGGGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCCCTCTGACCAGCCACCCAGC
GGCGCCTACGCTGATGCCTGCTGTCAACTATGCCCCCTTGGA

AA236353

GCCCGCTGGAGTCTGCAAATGCTGCCTTCCAAAAGGGACCTGGTGGGTTCTGTTCTCCCT
GGCAACACATCTGGCTGTTCCAGCCACCAGCGAGACGCCAAGACTGGTAACTGTCCACAG
GCAACAGGGAGTCATGTCGCGGTTCTCCAGGAGCACCCAGAGTCTGCACCACAGACACGT
CCAGGTAACTGGCCATAGGTNGGTAGGTTCCCGGATATCCCGGATAGAAGGCAAACTCAG
TGGGGCGGCTGGGGTACTCTTCCCCGGCCGTGGGAGTCTCCGCGGGGTACGCGCACAGGG
TGGCTGCCTGGGCACAGGGTTTCAGCGAGCTCCGGGACACTCGGCAGGAGTAGTACCCGC
CTCCAAAGTAACCATAAGGCA

AA588193

AACTGCTCGCCACCGACCCCACTACTCGCCACCGACCCGCTGCTCGGAGCTTCGGTTCTG
CGGGTTGTCCAGACTTCAGGCCTGTGCGCTCAATCGTGGAGAATGCGCCGGCAGCCCCCA
CCCCCAGCCTAAGGTGCAGGAAGGACCAGCACGAACCCGCTGGCTTTGCTGCGCGGCCAG
GAGATGAGTCCCACCGGGCACTGAGCCCAGGTACAGGACATCAGAGAATGAACACAGAGG
CAGAGGCCCTCATGTCCCTCTCAGAGTCCCGGCTCTGCAAAGAGCCCGTCTGTCTCCAGC
TTCCAGAATTCCGCACTGTGAATCTGTCTACGTGGACTGGGAAAACAGGGTTGGCACCAC
TCTGCCACTCCGTTTGTGCCTGGGAAGGGCTAAGTATGCAAGGCT

AI821103

GATCCCTTTGCAGGGAAGCTTTCTCTCAGACCCCCTTCCATTACACCTCTCACCCTGGTA
ACAGCAGGAAGACTGAGGAGAGGGGAACGGGCAGATTCGTTGTGTGGCTGTGATGTCCGT
TTAGCATTTTTCTCAGCTGACAGCTGGGTAGGTGGACAATTGTAGAGGCTGTCTCTTCCT
CCCTCCTTGTCCACCCCATAGGGTGTACCCACTGGTCTTGGAAGCACCCATCCTTAATAC
GATGATTTTTCTGTCGTGTGAAAATGAAGCCAGCAGGCTGCCCCTAGTCAGTCCTTCCTT
CCAGAGAAAAAGAGATTTGAGAAAGTGA

AI821851

TTTTTTTTTTTTTTTTTTTCTTTTTCACTTTCTCAAATCTCTTTTTCTCTGGAAGGAAG
GACTGACTAGGGGCAGCCTGCTGGCTTCATTTTCACACGACAAAAAAATCATCGTATTAA
GGATGGGTGCTTCCAAAACCAGTGGGTACACCCTATGGGGGGGACAAGGAGGGAGGAAGA
GACAGCCTCTACAATTGTCCACCTACCCAGCTGTCAGCTGAGAAAAATGCTAAACGGACA
TCACAGCCACACAACGAATCTGCCCGTTCCCCTCTCCTCAGTCTTCCTGCTGTTACCAGG
GTGAGAGGTGTAATGGAAGG

AA635855

TTTTTTTTTTTTTTTTTTTCTTTTTCACTTTCCCAAATCTCTTTTTCTCTGGAAGGAAG
GACTGACTAGGGGCAGCCTGCTGGCTTCATTTTCACACGACAGAAAAATCATCGTATTAA
GGATGGGTGCTTCCAAGACCAGTGGGTACACCCTATGGGGTGGACACAGGAGGGAGGAAG
AGACAGCCTCTACAATTGTCCACCTACCCAGCTGTCAGCTGAGAAAAATGCTAAACGGAC
ATCACAGCCACACAACGAATCTGCCCGTTCCCCTCTCCTCAGTCTTCCTGCTGTTACCAG
GGTGAGAGGTGTAATGGAAGG

AI420753

GCGGCCGCGGCCCACCACCAACTGCTCGCCACCGACCCCACTACTCGCCACCGACCCGCT
GCTCGGAGCTTCGGTTCTGCGGGTTGTCCAGACTTCAGGCCTGTGCGCTCAATCTTGGAG
AATGCGCCGGCAGGCCCCCCACCCCCAGCCTAAGGTGCAGGAAGGACCAGCACGAACCCG
CTGGCTTTGCTGCGCGGCCAGGAGATGAGTCCCACCGGGCACTGAGCCCAGGTACAGGAC

ATCAGAGAATGAACACAGAGGCAGAGGCCCTCATGTCCCTCTCAGAGTCCCGGCTCTGCA
AAGAGCCCGTCTGTCTCCAGCTTCCAGAATTCCGCACTGTGAATCTGTCTACGT

BG180547

CACGCGTCGATCCCAGTGAAGTAGATGTTTGTAGCCTTGCATACTTAGTCCTTCCCAGGC
ACAAACGGAGTGGCAGAGTGGTGCCAACCCTGTTTTCCCAGTCCACGTAGACAGATTCAC
AGTGCGGAATTCTGGAAGCTGGAGACAGACGGGCTCTTTGCAGAGCCGGGACTCTGAGAG
GGACATGAGGGCCTCTGCCTCTGTGTTCATTCTCTGATGTCCTGTACCTGGGCTCAGTGC
CCGGTGGGACTCATCTCCTGGCCGCGCAGCAAAGCCAGCGGGTTCGTGCTGGTCCTTCCT
GCACCTTAGGCTGGGGGTGGGGGGCCTGCCGGCGCATTCTCCACGATTGAGCGCACAGGC
CTGAAGTCTGGACAACCCGCAGAACCGAAGCTCCGAGCAGCGGGTCGGTGGCGAGTAGTG
GGGTCGGTGGCGAGCAGTTGGTGGTGGG

AA468306

TCGACCTCGCCAAGGTGAAGAACAACGCTACCCCTTAAGAGATCTCCTTGCCTGGGTGGG
AGGAGCGAAAGTGGGGGTGTCCTGGGGAGACCAGGAACCTGCCAAGCCCAGGCTGGGGCC
AAGGACTCTGCTGAGAGGCCCCTAGAGACAACACCCTTCCCAGGCCACTGGCTGCTGGAC
TGTTCCTCAGGAGCGGCCTGGGTACCCAGTATGTGCAGGGAGA

AA468232

TTTTTTACTGGTTATCGTGGTTATTGCCACTGTCAGGATGAATGATTATGACTGGGCCAG
GTTCTTTGGGAACCCTGGTGGAGTGGGCTGTCACATGGGGTTCCGTCTCCCTGCACATAC
TGGGTACCCAGGCCGCTCCTGAGGAACAGTCCAGCAG

CB050115

GGCCCACCACCAACTGCTCGCCACCGACCCCACTACTCGCCACCGACCCGCTGCTCGGAG
CTTCGGTTCTGCGGGTTGTCCAGACTTCAGGCCTGTGCGCTCAATCGTGGAGAATGCGCC
GGCAGGCCCCCCACCCCCAGCCTAAGGTGCAGGAAGGACCAGCACGAACCCGCTGGCTTT
GCTGCGCGGCCAGGAGATGAGTCCCACCGGGCACTGAGCCCAGGTACAGGACATCAGAGA
ATGAACACAGAGGCAGAGGCCCTCATGTCCCTCTCAGAGTCCCGGCTCTGCAAAGAGCCC
GTCTGTCTCCAGCTTCCAGAATTCCGCACTGTGAACCTCGTGCC

CB050116

GGCACGAGGTTCACAGTGCGGAATTCTGGAAGCTGGAGACAGACGGGCTCTTTGCAGAGC
CGGGACTCTGAGAGGGACATGAGGGCCTCTGCCTCTGTGTTCATTCTCTGATGTCCTGTA
CCTGGGCTCAGTGCCCGGTGGGACTCATCTCCTGGCCGCGCAGCAAAGCCAGCGGGTTCG
TGCTGGTCCTTCCTGCACCTTAGGCTGGGGGTGGGGGGCCTGCCGGCGCATTCTCCACGA
TTGAGCGCACAGGCCTGAAGTCTGGACAACCCGCAGAACCGAAGCTCCGAGCAGCGGGTC
GGTGGCGAGTAGTGGGGTCGGTGGCGAGCAGTTGGTGGTGGGCC

AA661819

GCTGCTCGGAGCTTCGGTTCTGCGGGTTGTCCAGACTTCAGGCCTGTGCGCTCAATCGTG
GAGAATGCGCCGGCAGCCCCCACCCCCAGCCTAAGGTGCAGGAAGGACCAGCACGAACCC
GCTGGCTTTGCTGCGCGGCCAGGAGATGAGTCCCACCGGCACTGAGCCAGGTACAGGACA
TCAGAGAATGAACACAGAGGCAGAGGCCTCATGTCCCTCTCAGAGTCCCGGCTCTGCAAA
GAGCCGTACTGTCTCCAGCTTCCAGAATTCCGCACTGTGAATCTGTCTACGTGGACTGGG
AAAAC

CF146837

CACGAGGATTTTCTATCTAGAGCTCTGTAGAGCACTTTAGAAACCGCTTTCATGAATTGA
GCTAATTATGAATAAATTTGGAAGGCGATCCCTTTGCAGGGAAGCTTTCTCTCAGACCCC
CTTCCATTACACCTCTCACCCTGGTAACAGCAGGAAGACTGAGGAGAGGGGAACGGGCAG
ATTCGTTGTGTGGCTGTGATGTCCGTTTAGCATTTTTCTCAGCTGACAGCTGGGTAGGTG
GACAATTGTAGAGGCTGTCTCTTCCTCCCTCCTTGTCCACCCCATAGGGTGTACCCACTG
GTCTTGGAAACACCCATCCTTAATACGATGATTTTTCTGTCGTGTGAAAATGAAGCCAGC
AGGCTGCCCCTAGTCAGTCCTTCCTTCCAGAGAAAAAGAGATTTGAGAAAGTGCCTGGGT
AATTCACCATTAATTTCCTCCCCCAAACTCTCTGAGTCTTCCCTTAATATTTCTGGTGGT
TCTGACCAAAGCAGGTCATGGTTTGTTGAGCATTTGGGATCCCAGTGAAGTAGATGTTTG
TAGCCTTGCATACTTAGCCCTTCCCAGGCACAAACGGAGTGGCAGAGTGGTGCCAACCCT
GTTTTCCCAGTCCACGTAGACAGATTCACAGTGCGGAATTCTGGAAGCTGGAGACAGACG
GGCTCTTTGCAGAGCCGGGACTCTGAG

CF146763

CACGAGGATTTTCTATNCTAGAGCTCTGGTAGAGCACTTTANAAACCGCTTTCATGAATT
GAGCTAATTATGAATAAATTTGGAAGGCGATCCCTTTGCAGGGAAGCTTTCTCTCAGACC
CCCTTCCATTACACCTCTCACCCTGGTAACAGCAGGAAGACTGAGGAGAGGGGAACGGGC
AGATTCGTTGTGTGGCTGTGATGTCCGTTTAGCATTTTTCTCAGCTGACAGCTGGGTAGG
TGGACAATTGTAGAGGCTGTCTCTTCCTCCCTCCTTGTCCACCCCATAGGGTGTACCCAC
TGGTCTTGGAAACACCCATCCTTAATACGATGATTTTTCTGTCGTGTGAAAATGAAGCCA
GCAGGCTGCCCCTAGTCAGTCCTTCCTTCCAGAGAAAAAGAGATTGAGAAAGTGCCTGGG
TAATTCACCATTAATTTCCTCCCCCAAACTCTCTGAGTCTTCCCTTAATATTTCTGGTGG
TTCTGACCAAAGCAGGTCATGGTTTGTTGAGCATTTGGGATCCCAGTGAAGTAGATGTTT
GTAGCCTTGCATACTTAGCCCTTCCCAGGCACAAACGGAGTGGCAGAGTGGTGCCAACCC
TGTTTTCCCAGTCCACGTAGACAGATTCACAGTGCGGAATTCTGGAAGCTGGAGACAGAC
GGGCTCTTTGCAGAGCCGGGACTCTGA

CF144902

CACGAGGGAAGCCAGCAGGCTGCCCCTAGTCAGTCCTTCCTTCCAGAGAAAAAGAGATTT
GAGAAAGTGCCTGGGTAATTCACCATTAATTTCCTCCCCCAAACTCTCTGAGTCTTCCCT

TAATATTTCTGGTGGTTCTGACCAAAGCAGGTCATGGTTTGTTGAGCATTTGGGATCCCA
GTGAAGTAGATGTTTGTAGCCTTGCATACTTAGCCCTTCCCAGGCACAAACGGAGTGGCA
GAGTGGTGCCAACCCTGTTTTCCCAGTCCACGTAGACAGATTCACAGTGCGGAATTCTGG
AAGCTGGAGACAGACGGGCTCTTTGCAGAGCCGGGACTCTGAGAGGGACATGAGGGCCTC
TGCCTCTGTGTTCATTCTCTGATGTCCTGTACCTGGGCTCAGTGCCCGGTGGGACTCATC
TCCTGGGCGCGCAGCAAAGCCAGCGGGTTCGTGCTGGTCCTTCCTGCACCTTA

CF141511.1

CACGAGGCCTGGTAACAGCAGGAAGACTGAGGAGAGGGGAACGGGCAGATTCGTTGTGTG
GCTGTGATGTCCGTTTAGCATTTTTCTCAGCTGACAGCTGGGTAGGTGGACAATTGTAGA
GGCTGTCTCTTCCTCCCTCCTTGTCCACCCCATAGGGTGTACCCACTGGTCTTGGAAACA
CCCATCCTTAATACGATGATTTTTCTGTCGTGTGAAAATGAAGCCAGCAGGCTGCCCCTA
GTCAGTCCTTCCTTCCAGAGAAAAAGAGATTTGAGAAAGTGCCTGGGTAATTCACCATTA
ATTTCCTCCCCCAAACTCTCTGAGTCTTCCCTTAATATTTCTGGTGGTTCTGACCAAAGC
AGGTCATGGTTTGTTGAGCATTTGGGATCCCAGTGAAGTAGATGTTTGTAGCCTTGCATA
CTTAGCCCTTCCCAGGCACAAACGGAGTGGCAGAGTGGTGCCAACCCTGTTTTCCCAGTC
CACGTAGACAGATTCACAGTGCGGAATTCTGGAA

CF139563.1

CACGAGGTCTTCCCTTAATATTTCTGGTGGTTCTGACCAAAGCAGGTCATGGTTTGTTGA
GCATTTGGGATCCCAGTGAAGTAGATGTTTGTAGCCTTGCATACTTAGCCCTTCCCAGGC
ACAAACGGAGTGGCAGAGTGGTGCCAACCCTGTTTTCCCAGTCCACGTAGACAGATTCAC
AGTGCGGAATTCTGGAAGCTGGAGACAGACGGGCTCTTTGCAGAGCCGGGACTCTGAGAG
GGACATGAGGGCCTCTGCCTCTGTGTTCATTCTCTGATGTCCTGTACCTGGGCTCAGTGC
CCGGTGGGACTCATCTCCTGGCCGCGCAGCAAAGCCAGCGGGTTCGTGCTGGTCCTTCCT
GCACCTTAGGCTGGGGGTGGGGGGCCTGCCGGCGCATTCTCCACGATTGAGCGCACAGGC
CTGAAGTCTGGACAACCCGCAGAACCGAAGCTCCGAGCAGCGGGTCGGTGGCGAGTA

CF139372

CACGAGGATTTCTGGTGGTTCTGACCAAAGCAGGTCATGGTTTGTTGAGCATTTGGGATC
CCAGTGAAGTAGATGTTTGTAGCCTTGCATACTTAGCCCTTCCCAGGCACAAACGGAGTG
GCAGAGTGGTGCCAACCCTGTTTTCCCAGTCCACGTAGACAGATTCACAGTGCGGAATTC
TGGAAGCTGGAGACAGACGGGCTCTTTGCAGAGCCGGGACTCTGAGAGGGACATGAGGGC
CTCTGCCTCTGTGTTCATTCTCTGATGTCCTGTACCTGGGCTCAGTGCCCGGTGGGACTC
ATCTCCTGGCCGCGCAGCAAAGCCAGCGGGTTCGTGCTGGTCCTTCCTGCACCTT

CF139319

CACGAGGAAGGCGATCCCTTTGCAGGGAAGCTTTCTCTCAGACCCCCTTCCATTACACCT
CTCACCCTGGTAACAGCAGGAAGACTGAGGAGAGGGGAACGGGCAGATTCGTTGTGTGGC
TGTGATGTCCGTTTAGCATTTTTCTCAGCTGACAGCTGGGTAGGTGGACAATTGTAGAGG

CTGTCTCTTCCTCCCTCCTTGTCCACCCCATAGGGTGTACCCACTGGTCTTGGAAACACC
CATCCTTAATACGATGATTTTTCTGTCGTGTGAAAATGAAGCCAGCAGGCTGCCCCTAGT
CAGTCCTTCCTTCCAGAGAAAAAGAGATTTGAGAAAGTGCCTGGGTAATTCACCATTAAT
TTCCTCCCCCAAACTCTCTGAGTCTTCCCTTAATATTTCTGGTGGTTCTGACCAAAGCAG
GTCATGGTTTGTTGAGCATTTGGGATCCCAGTGAAGTAGATGTTTGTAGCCTTGCATACT
TAGCCCTTCC

CF139275

CACGAGGTGGATTCCCCCGGCCTGGGTGGGGAGAGCGAGCTGGGTGCCCCCTAGATTCCC
CGCCCCCGCACCTCATGAGCCGACCCTCGGCTCCATGGAGCCCGGCAATTATGCCACCTT
GGATGGAGCCAAGGATATCGAAGGCTTGCTGGGAGCGGGAGGGGGGCGGAATCTGGTCGC
CCACTCCCCTCTGAGCAGCCACCCAGCGGCGCCTACGCTGATGCCTGCTGTCAACTATGC
CCCCTTGGATCTGCCAGGCTCGGCGGAGCCGCCAAAGCAATGCCACCCATGCCCTGGGGT
GCCCCAGGGGACGTCCCCAGCTCCCGTGCCTTATGGTTACTTTGGAGGCGGGTACTACTC
CTGCCGAGTGTCGCGGAGCTCGCTGAAACCCTGTGCCCAGGCA

CF122893

CACGAGGATTTTCTATCTAGAGCTCTGTAGAGCACTTTAGAAACCGCTTTCATGAATTGA
GCTAATTATGAATAAATTTGGAAGGCGATCCCTTTGCAGGGAAGCTTTCTCTCAGACCCC
CTTCCATTACACCTCTCACCCTGGTAACAGCAGGAAGACTGAGGAGAGGGGAACGGGCAG
ATTCGTTGTGTGGCTGTGATGTCCGTTTAGCATTTTTCTCAGCTGACAGCTGGGTAGGTG
GACAATTGTAGAGGCTGTCTCTTCCTCCCTCCTTGTCCACCCCATAGGGTGTACCCACTG
GTCTTGGAAACACCCATCCTTAATACGATGATTTTTCTGTCGTGTGAAAATGAAGCCAGC
AGGCTGCCCCTAGTCAGTCCTTCCTTCCAGAGAAAAAGAGATTTGAGAAAGTGCCTGGGT
AATTCACCATTAATTTCCTCCCCCAAACTCTCTGAGTCTTCCCTTAATATTTCTGGTGGT
TCTGACCAAAGCAGGTCATGGTTTGTTGAGCATTTGGGATCCCAGTGAAGTANATGTTTG
TAGCCTTGCATACTTAGCCCTT

AI972423

CATTTTCACACGACTGTAAAATCATCGTATTAAGGATGGGTGCTTCCAAGACCAGTGGGT
ACACCCTATGGGGTGGACAAGGAGGGAGGAAGAGACAGCCTCTACAATTGTCCACCTACC
CAGCTGTCAGCTGAGAAAAATGCTAAACGGACATCACAGCCACACAACGAATCTGCCCGT
TCCCCTCTCCTCAGTCTTCCTGCTGTTACCAGGGTGAGAGGTGTAATGGAAGGGGGTCTG
AGAGAAAGCTTCCCTGCAAAGGGATCGCCTTCCAAATTTATTCATAATTAGCTCAATTCA
TGAAAGCGGTTTCTAAAGTGCTCTACAGAGCTCTAGATAGAAAATATGAGGCTAACGATC
ATGGCAGCTAGTACTGGTTATCGTGATTATTGCCACTGTCAGGATGAATGATTATGACTG
GGCCAGGTTCTTTGGGAACCCTGGTGGAGTGGGCTGTCACATG

AI918975

TGCAGCTAGTACTGGTTATCGTGATTATTGCCACTGTCAGGATGAATGATTATGACTGGG

CCAGGTTCTTTGGGAACCCTGGTGGAGTGGGCTGTCACATGGGGTTCCGTCTCCCTGCAC
ATACTGGGTACCCAGGCCGCTCCTGAGGAACAGTCCAGCACAGGGTTTCAGCGAGCTCCG
GGACACTCGGCCTCGTGC

AI826991

TTTTTTTTTTTTTTTTTTTCTTTTTCACTTTCTCAAATCTCTTTTTCTCTGGAAGGAAG
GACTGACTAGGGGCAGCCTGCTGGCTTCATTTTCACACCACAAAAAAATCATCGTATTAA
GGATGGGTGCTTCCAAAACCAGTGGGTACACCCTATGGGGTGGACAAGGAGGGAGGAAAA
AACAGCCTCTACAATTGTCCACCTACCCAGCTGTCAGCTGAAAAAAATGCTAAACGGACA
TCACAGCCACACAACGAATCTGCCCGTTCCCCTCTCCTCAGTCTTCCTGCTGTTACCAGG
GTGAAAGGTGTAATGGAAGG

AI686312

ACCGACCCCACTACTTGCCACCGACCCGCTGCTCGGAGCTTCGGTTCTGCGGGTTGTCCA
GACTTCAGGCCTGTGCGCTCAATCGTGGAGAATGCGCCGGCAGGCCCCCCACCCCCAGCC
TAAGGTGCAGGAAGGACCAGCACGAACCCGCTGGCTTTGCTGCGCGGCCAGGAGATGAGT
CCCACCGGGCACTGAGCCCAGGTACAGGACATCAGAGAATGAACACAGAGGCAGAGGCCC
TCATGTCCCTCTCAGAGTCCCGGCTCTGCAAAGAGCCCGTCTGTCTCCAGCTTCCAGAAT
TCCGCACTGTGAATCTGTCTACGTGGACTGGGAAAACAGGGTTGGCACCACTCTGCCACT
CCGTTTGTGCCTGGGAAGGGCTAAGTATGCAAGGCTACAAACATCTACTTCACTGGGATC
C

AI655923

TTTTTTTTTTTTTTTCCCTGCAAAGGGATCGCCTTCCAAATTTATTCATAATTAGCTCAA
TTCATGAAAGCGGTTTCTAAAGTGCTCTACAGAGCTCTAGATAGAAAATATGAGGCTAAC
GATCATGGCAGCTAGTACTGGTTATCGTGATTATTGCCACTGTCAGGATGAATGATTATG
ACTGGGCCAGGTTCTTTGGGAACCCTGGTGGAGTGGGCTGTCACATGGGGTTCCGTCTCC
CTGCACATACTGGGTACCCAGGCCGCTCCTGA

CF146922

CACGAGGCGACTTGCGAGCTGGGAGCGATTTAAAACGCTTTGGATTCCCCGGCCTGGGTG
GGGAGAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCGCACCTCATGAGCCGACCCTC
GGCTCCATGGAGCCCGGCAATTATGCCACCTTGGATGGAGCCAAGGATATCGAAGGCTTG
CTGGGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCCCTCTGACCAGCCACCCAGCG
GCGCCTACGCTGATGCCTGCTGTCAACTATGCCCCCTTGGATCTGCCAGGCTCGGCGGAG
CCGCCAAAGCAATGCCACCCATGCCCTGGGGTGCCCCAGGGGACGTCCCCAGCTCCCGTG
CCTTATGGTTACTTTGGAGGCGGGTACTACTCCTGCCGAGTGTCCCGGAGCTCGCTGAAA
CCCTGTGCCCAGGCAGCCACCCTGGCCGCGTACCCCGCGGAGACTCCCACGGCCGGGGAA
GAGTACCCCAGCCGCCCCACTGAGTTTGCCTTCTATCCGGGATATCCGGGAACCTACCAG
CCTATGGCCAGTTACCTGGACGTGTCTGTGGTGCAGACTCTGGGTGCTCCTGGAGAACGC

GACATGACTCCCTGTTGCCTGTGGACAGTTACCAGTCTTGGGCTCTCGCTGGTGGCTGGA
ACAGCCAGATGTGTTGCCA

BF476369

GCGGCCGCGGCCCACCACCAACTGCTCGCCATTCGACCCCACTACTCGCCACCGACCCGC
TGCTCGGAGCTTCGGTTCTGCGGGTTGTCCAGACTTCAGGCCTGTGCGCTCAATCGTGGA
GAATGCGCCGGCAGGCCCCCCACCCCCAGCCTAAGGTGCAGGAAGGACCAGCACGAACCC
GCTGGCTTTGCTGCGCGGCCAGGAGATGAGTCCCACCGGGCACTGAGCCCAGGTACAGGA
CATCAGAGAATGAACACAGAGGCAGAGGCCCTCATGTCCCTCTCAGAGTCCCGGCTCTGC
AAAGAGCCCGTCTGTCTCCAGCTTCCAGAATTCCGCACTGTGAATCTGTCTACGTGGACT
GGGAAAACAGGGTTGGCACCACTCTGCCACTCC

BF057410

GCGGCCGCGGCCCACCACCAACTGCTCGCCACCGACCCCACTACTCGCCACCGACCCGCT
GCTCGGAGCTTCGGTTCTGCGGGTTGTCCAGACTTCAGGCCTGTGCGCTCAATCGTGGAG
AATGCGCCGGCAGGCCCCCCACCCCCAGCCTAAGGTGCAGGAAGGACCAGCACGAACCCG
CTGGCTTTGCTGCGCGGCCAGGAGATGAGTCCCACCGGGCACTGAGCCCAGGTACAGGAC
ATCAGAGAATGAACACAGAGGCAGAGGCCCTCATGTCCCTCTCAGAGTCCCGGCTCTGCA
AAGAGCCCGTCTGTCTCCAGCTTCCAGAATTCCGCACTGTGAATCTGTCTACGTGGACTG
GGAAAACAGGGTTGGCACCACTCTGCCACTCCGTTTGTGCCTGGGAAGGGCTAAGTATGC
AAGGCTACAAACATCTACTTCACTGGGATCCCAAATGCTCAACAAACCATGACCTGCTNT
GGTCAGAACCACCAGAAATATTAA

BE645544

GCGGCCGCGGCCCACCACCAACTGCTCGCCACCGACCCCACTACTCGCCACCGACCCGCT
GCTCGGAGCTTCGGTTCTGCGGGTTGTCCAGACTTCAGGCCTGTGCGCTCAATCGTGGAG
AATGCGCCGGCAGGCCCCCCACCCCCAGCCTAAGGTGCAGGAAGGACCAGCACGAACCCG
CTGGCTTTGCTGCGCGGCCAGGAGATGAGTCCCACCGGGCACTGAGCCCAGGTACAGGAC
ATCAGAGAATGAACACAGAGGCAGAGGCCCTCATGTCCCTCTCAGAGTCCCGGCTCTGCA
AAGAGCCCGTCTGTCTCCAGCTTCCAGAATTCCGCACTGTGAATCTGTCTACGTGGACTG
GGAAAACAGGGTTGGCACCACTCTGCCACTCCGTTTGTGCCTGGGAAGGGCTAAGTATGC
AAGGCTACAAACATCTACTTCACTGGGATCC

BE645408

TCCTCCCTCTAAGAAAGGCGCAAGCGTCAAGAGGGTGCTGCCCGCTGGTTTCTGCAAATG
CTGCCTTCCAAAAAGGACCTGGTGGGTTCTGTTCTCCCTGGCAACACATCTGGCTGTTCC
AGCCACCAGCGAGAGCCCAAGACTGGTAACTGTCCACAGGCAACAGGGAGTCATGTCGCG
GTTCTCCAGGAGCACCCAGAGTCTGCACCACAGACACGT

BE388501

TTAATACGATGATTTTTCTGTCGTGTGAAAATGAAGCCAGCAGGCTGCCCCTAGTCAGTC
CTTCCTTCCAGAGAAAAAGAGATTTGAGAAAGTGCCTGGGTAATTCACCATTAATTTCCT
CCCCCAAACTCTCTGAGTCTTCCCTTAATATTTCTGGTGGTTCTGACCAAAGCAGGTCAT
GGTTTGTTGAGCATTTGGGATCCCAGTGAAGTAGATGTTTGTAGCCTTGCATACTTAGCC
CTTCCCAGGCACAAACGGAGTGGCAGAGTGGTGCCAACCCTGTTTTCCCAGTCCACGTAG
ACAGATTCACAGTGCGGAATTCTGGAAGCTGGAGACAGACGGGCTCTTTGCAGAGCCGGG
ACTCTGAGAGGGACATGAGGGCCTCTGCCTCTGTGTTCATTCTCTGATGTCCTGTACCTG
GGCTCAGTGCCCGGTGGGACTCATCTCCTGGCCGCGCAGCAAAGCCAGCGGGTTCGTGCT
GGTCCTTCCTGCACCTTAGGCTGGGGGTGGGGGGCCTGCCGGCGCATTCTCCACGATTGA
GCGCACAGGCCTGAAGTCTGGACAACCCGCAGAACCGAAGCTCCGAGCAGCGGGTCGGTG
GCGAGTAGTGGGGGTCGGTGGCGAACAAGTGGTGGTGGGCCGGGGCCGCATAACTCGAGG
ACTTTCCTCCCGGAGCAGTCCCTAAAAACCCGGGGGCGC

CF147366

GACGAGGACAATTGTAGAGGCTGTCTCTTCCTCCCTCCTTGTCACCCCATAGGGTGTACC
ACTGGTCTTGGAAGCACCCATCCTTAATACGATGATTTTTCTGTCGTGTGAAAATGAAGC
CAGCAGGCTGCCCCTAGTCAGTCCTTCCTTCCAGAGAAAAAGAGATTTGAGAAAGTGCCT
GGGTAATTCACCATTAATTTCCTCCCCCAAACTCTCTGAGTCTTCCCTTAATATTTCTGG
TGGTTCTGACCAAAGCAGGTCATGGTTTGTTGAGCATTTGGGATCCCAGTGAAGTAGATG
TTTGTAGCCTTGCATACTTAGCCCTTCCCAGGCACAAACGGAGTGGCAGAGTGGTGCCAA
CCCTGTTTTCCCAGTCCACGTAGACAGATTCACAGTGCGGAATTCTGGAAGCTGGAGACA
GACGGGCTCTTTGCAGAGCCGGGACTCTGAGAGGGACATGAGGGCCTCTGCCTCTGTGTT
CATTCTCTGATGTCCTGTACCTGGGCTCAGTGCCCGGTGGGACTCATCTCCTGGCCGCGC
AGCAAAGCCAGCGGGTTCGTGCTGGTCCTTCCTGC

CF147143

CACGAGGCGACTTGCGAGCTGGGAGCGATTTAAAACGCTTTGGATTCCCCCGGCCTGGGT
GGGGAGAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCGCACCTCATGAGCCGACCCT
CGGCTCCATGGAGCCCGGCAATTATGCCACCTTGGATGGAGCCAAGGATATCGAAGGCTT
GCTGGGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCCCTCTGACCAGCCACCCAGC
GGCGCCTACGCTGATGCCTGCTGTCAACTATGCCCCCTTGGATCTGCCAGGCTCGGCGGA
GCCGCCAAAGCAATGCCACCCATGCCCTGGGGTGCCCCAGGGGACGTCCCCAGCTCCCGT
GCCTTATGGTTACTTTGGAGGCGGGTACTACTCCTGCCGAGTGTCCCGGAGCTCGCTGAA
ACCCTGTGCCCAGGCAGCCACCCTGGCCGCGTACCCCGCGGAGACTCCCACGGCCGGGGA
AGAGTACCCAGCCGCCCCACTGAGTTTGCCTTCTATCCGGGATATCCGGGAACCTACCAG
CCTATGGCCAGTTACCTGGACGTGTCTGTGGTGCAGACTCTGGGTGCTCCTGGAGAACGC
GACATGACTCCCTGTTGCCTGTGGACAGTTACCAATCTTGGGCTCTCGCTGGTGGCTGGA
ACAGCCAGATGTGTTGCCAGGGAG

BT007410

atggagcccg gcaattatgc caccttggat ggagccaagg atatcgaagg cttgctggga

gcgggagggg ggcggaatct ggtcgcccac tcccctctga ccagccaccc agcggcgcct
acgctgatgc ctgctgtcaa ctatgccccc ttggatctgc caggctcggc ggagccgcca
aagcaatgcc acccatgccc tggggtgccc caggggacgt ccccagctcc cgtgccttat
ggttactttg gaggcgggta ctactcctgc cgagtgtccc ggagctcgct gaaaccctgt
gcccaggcag ccaccctggc cgcgtacccc gcggagactc ccacggccgg ggaagagtac
cccagccgcc ccactgagtt tgccttctat ccgggatatc cgggaaccta ccagcctatg
gccagttacc tggacgtgtc tgtggtgcag actctgggtg ctcctggaga accgcgacat
gactccctgt tgcctgtgga cagttaccag tcttgggctc tcgctggtgg ctggaacagc
cagatgtgtt gccagggaga acagaaccca ccaggtccct tttggaaggc agcatttgca
gactccagcg ggcagcaccc tcctgacgcc tgcgcctttc gtcgcggccg caagaaacgc
attccgtaca gcaaggggca gttgcgggag ctggagcggg agtatgcggc taacaagttc
atcaccaagg acaagaggcg caagatctcg gcagccacca gcctctcgga gcgccagatt
accatctggt ttcagaaccg ccgggtcaaa gagaagaagg ttctcgccaa ggtgaagaac
agcgctaccc cttag

BC007092

ggattccccc ggcctgggtg gggagagcga gctgggtgcc ccctagattc cccgcccccg
cacctcatga gccgaccctc ggctccatgg agcccggcaa ttatgccacc ttggatggag
ccaaggatat cgaaggcttg ctgggagcgg gagggggggcg gaatctggtc gcccactccc
ctctgaccag ccacccagcg gcgcctacgc tgatgcctgc tgtcaactat gccccccttgg
atctgccagg ctcggcggag ccgccaaagc aatgccaccc atgccctggg gtgccccagg
ggacgtcccc agctcccgtg ccttatggtt actttggagg cgggtactac tcctgccgag
tgtcccggag ctcgctgaaa ccctgtgccc aggcagccac cctggccgcg taccccgcgg
agactcccac ggccggggaa gagtacccca gccgccccac tgagtttgcc ttctatccgg
gatatccggg aacctaccag cctatggcca gttacctgga cgtgtctgtg gtgcagactc
tgggtgctcc tggagaaccg cgacatgact ccctgttgcc tgtggacagt taccagtctt
gggctctcgc tggtggctgg aacagccaga tgtgttgcca gggagaacag aacccaccag
gtccctttg gaaggcagca tttgcagact ccagcgggca gcaccctcct gacgcctgcg
cctttcgtcg cggccgcaag aaacgcattc cgtacagcaa ggggcagttg cgggagctgg
agcgggagta tgcggctaac aagttcatca ccaaggacaa gaggcgcaag atctcggcag
ccaccagcct ctcggagcgc cagattacca tctggtttca gaaccgccgg gtcaaagaga
agaaggttct cgccaaggtg aagaacagcg ctacccctta agagatctcc ttgcctgggt
gggaggagcg aaagtggggg tgtcctgggg agaccaggaa cctgccaagc ccaggctggg
gccaaggact ctgctgagag gcccctagag acaacaccct tcccaggcca ctggctgctg
gactgttcct caggagcggc ctgggtaccc agtatgtgca gggagacgga accccatgtg
acagcccact ccaccagggt tcccaaagaa cctggcccag tcataatcat tcatcctgac
agtggcaata atcacgataa ccagtactag ctgccatgat cgttagcctc atattttcta
tctagagctc tgtagagcac tttagaaacc gctttcatga attgagctaa ttatgaataa
atttggaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa

U57052

cgggtgcccc ctagattccc cgcccccgca cctcatgagc cgaccctcgg ctccatggag

cccggcaatt atgccacctt ggatggagcc aaggatatcg aaggcttgct gggagcggga
gggggggcgga atctggtcgc ccactcccct ctgaccagcc acccagcggc gcctacgctg
atgcctgctg tcaactatgc ccccttggat ctgccaggct cggcggagcc gccaaagcaa
tgccacccat gccctggggt gccccagggg acgtccccag ctcccgtgcc ttatggttac
tttggaggcg ggtactactc ctgccgagtg tcccggagct cgctgaaacc ctgtgcccag
gcagccaccc tggccgcgta ccccgcggag actcccacgg ccggggaaga gtaccccagc
cgccccactg agtttgcctt ctatccggga tatccgggaa cctaccacgc tatggccagt
tacctggacg tgtctgtggt gcagactctg ggtgctcctg gagaaccgcg acatgactcc
ctgttgcctg tggacagtta ccagtcttgg gctctcgctg gtggctggaa cagccagatg
tgttgccagg gagaacagaa cccaccaggt ccctttgga aggcagcatt tgcagactcc
agcgggcagc accctcctga cgcctccgcc tttcgtcgcg gccgcaagaa acgcattccg
tacagcaagg ggcagttgcg ggagctggag cgggagtatg cggctaacaa gttcatcacc
aaggacaaga ggcgcaagat ctcggcagcc accagcctct cggagcgcca gattaccatc
tggtttcaga accgccgggt caaagagaag aaggttctcg ccaaggtgaa gaacagcgct
acccttaag agatctcctt gcctgggtgg gaggagcgaa agtgggggtg tcctggggag
accaggaacc tgccaagccc aggctggggc caaggactct gctgagaggc ccctagagac
aacacc

U81599

tcctaatacg actcactata gggctcgagc ggccgcccgg gcaggtcgaa tgcaggcgac
ttgcgagctg ggagcgattt aaaacgcttt ggattccccc ggcctgggtg gggagagcga
gctgggtgcc ccctagattc cccgcccccg cacctcatga gccgaccctc ggctccatgg
agcccggcaa ttatgccacc ttggatggag ccaaggatat cgaaggcttg ctgggagcgg
gaggggggcg gaatctggtc gcccactccc ctctgaccag ccacccagcg gcgcctacgc
tgatgcctgc tgtcaactat gccccccttgg atctgccagg ctcggcggag ccgccaaagc
aatgccaccc atgccctggg gtgccccagg ggacgtcccc agctcccgtg ccttatggtt
actttggagg cgggtactac tcctgccgag tgtcccggag ctcgctgaaa ccctgtgccc
aggcagccac cctggccgcg taccccgcgg agactcccac ggccggggaa gagtacccca
gtcgccccac tgagtttgcc ttctatccgg gatatccggg aacctaccac gctatggcca
gttacctgga cgtgtctgtg gtgcagactc tgggtgctcc tggagaaccg cgacatgact
ccctgttgcc tgtggacagt taccagtctt gggctctcgc tggtggctgg aacagccaga
tgtgttgcca gggagaacag aacccaccag gtccctttg gaaggcagca tttgcagact
ccagcgggca gcaccctcct gacgcctgcg cctttcgtcg cggccgcaag aaacgcattc
cgtacagcaa ggggcagttg cgggagctgg agcgggagta tgcggctaac aagttcatca
ccaaggacaa gaggcgcaag atctcggcag ccaccagcct ctcggagcgc cagattacca
tctggtttca gaaccgccgg gtcaaagaga agaaggttct cgccaaggtg aagaacagcg
ctacccctta agagatctcc ttgcctgggt gggaggagcg aaagtggggg tgtcctgggg
agaccagaaa cctgccaagc ccaggctggg gccaaggact ctgctgagag gcccctagag
acaacaccct tcccaggcca ctggctgctg gactgttcct caggagcggc ctgggtaccc
agtatgtgca gggagacgga accccatgtg acaggcccac tccaccaggg ttcccaaaga
acctggccca gtcataatca ttcatcctca cagtggcaat aatcacgata accagt

CB120119

ATTTTTCTGTCGTGTGAAAATGAAGCCAGCAGGCTGCCCCTAGTCAGTCCTTCCTTCCAG
AGAAAAAGAGATTTGAGAAAGTGCCTGGGTAATTCACCATTAATTTCCTCCCCCAAACTC
TCTGAGTCTTCCCTTAATATTTCTGGTGGTTCTGACCAAAGCAGGTCATGGTTTGTTGAG
CATTTGGGATCCCAGTGAAGTAGATGTTTGTAGCCTTGCATACTTAGCCCTTCCCAGGCA
CAAACGGAGTGGCAGAGTGGTGCCAACCCTGTTTTCCCAGTCCACGTAGACAGATTCACA
GTGCGGAATTCTGGAAGCTGGAGACAGACGGGCTCTTTGCAGAGCCGGGACTCTGAGAGG
GACATGAGGGCCTCTGCCTCTGTGTTCATTCTCTGATGTCCTGTACCTGGGCTCAGTGCC
CGGTGGGACTCATCTCCTGGCTGCGCAGCAAAGCCAGCGGGTTCGTGCTGGTCCTTCCTG
CACCTTAGGCTGGGGGTGGGGGGCCT

CB125764

ATTTTTCTGTCGTGTGAAAATGAAGCCAGCAGGCTGCCCCTAGTCAGTCCTTCCTTCCAG
AGAAAAAGAGATTTGAGAAAGTGCCTGGGTAATTCACCATTAATTTCCTCCCCCAAACTC
TCTGAGTCTTCCCTTAATATTTCTGGTGGTTCTGACCAAAGCAGGTCATGGTTTGTTGAG
CATTTGGGATCCCAGTGAAGTAGATGTTTGTAGCCTTGCATACTTAGCCCTTCCCAGGCA
CAAACGGAGTGGCAGAGTGGTGCCAACCCTGTTTTCCCAGTCCACGTAGACAGATTCACA
GTGCGGAATTCTGGAAGCTGGAGACAGACGGGCTCTTTGCAGAGCCGGGACTCTGAGAGG
GACATGAGGGCCTCTGCCTCTGTGTTCATTCTCTGATGTCCTGTACCTGGGCTCAGTGCC
CGGTGGGACTCATCTCCTGGCTGCGCAGCAAAGCCAGCGGGTTCGTGCTGGTCCTTCCTG
CACCTTAGGCTGGGGGTGGGGGGGCCTGCCGGCGCATTCTCCACGATTGAGCGCACAGG
CCTGAAGTCTGGACAACCCGCAGAACCGAAGCTCCGAGCAGCGGGTCGGTGGCGAGT

AU098628

ATTTAAAACGCTTTGGATTCTTTCGTCCTGCGTGGGGAGAGCGAGCTGGGTGCCCCCTAG
ATTCCCCGCCCCCGCACCTCATGAGCCGACCCTCGGCTCCATGGAGCCCGGCACTTATGC
CACCTTGGATGGAGCCAAGGATATCGAAGGCTTGCTGGGAGCGGGAGGGGGGCGGAATCT
GGTCGCCCACTCCCCTCTGACCAGCCACCCAGCGGCGCCTACGCTGATGCCTGCTGTCAA
TTATGCCCCCTTGCATCTGCCAGGCTCGGCGGAGCCGCCAAAGCAATGCCACCCATGCCC

CB126130

ATTTTTCTGTCGTGTGAAAATGAAGCCAGCAGGCTGCCCCTAGTCAGTCCTTCCTTCCAG
AGAAAAAGAGATTTGAGAAAGTGCCTGGGTAATTCACCATTAATTTCCTCCCCCAAACTC
TCTGAGTCTTCCCTTAATATTTCTGGTGGTTCTGACCAAAGCAGGTCATGGTTTGTTGAG
CATTTGGGATCCCAGTGAAGTAGATGTTTGTAGCCTTGCATACTTAGCCCTTCCCAGGCA
CAAACGGAGTGGCAGAGTGGTGCCAACCCTGTTTTCCCAGTCCACGTAGACAGATTCACA
GTGCGGAATTCTGGAAGCTGGAGACAGACGGGCTCTTTGCAGAGCCGGGACTCTGAGAGG
GACATGAGGGCCTCTGCCTCTGTGTTCATTCTCTGATGTCCTGTACCTGGGCTCAGTGCC
CGGTGGGACTCATCTCCTGGCTGCGCAGCAAAGCCAGCGGGTTCGTGCTGGTCCTTCCTG
CACCTTAGGCTGGGGGTGGGGGGCCTGC

BI023924

AGGCCGCACCCAGTCTTAAGGTGCAGTGAAGGACAGCACGAACCCGCTGTGCTTTGCTGC
GCGGCAGGAGATGAGTCCCACCGGGCACTGAGCCCAGGTACAGGACATCAGAGAATGAAC
ACAGAGGCAGAGGCCCTCATGTCCCTCTCAGAGTCCCGGCTCTGCAAAGAGCCCGTCTGT
CTCCAGCTTCCAGAATTCCGCACTGTGAATCTGTCTACGTGGACTGNGAAAACAGGGTTG
GCACCACTCTGCCACTCCGTTTGTGCCTNGGGGCGGGCAGAGGG

BM767063.1

AAAAACGCTTTGGATTCCCCCGGCCTGGGTGGGGAGAGCGAGCTGGGTGCCCCCTAGATT
CCCCGCCCCCGCACCTCATGAGCCGACCCTCGGCTCCATGGAGCCCGGCAATTATGCCAC
CTTGGATGGAGCCAAGGATATCGAAGGCTTGCTGGGAGCGGGAGGGGGGCGGAATCTGGT
CGCCCACTCCCCTCTGACCAGCCACCCAGCGGCGCCTACGCTGATGCCTGCTGTCAACTA
TGCCCCCTTGGATCTGCCAGGCTCGGCGGAGCCGCCAAAGCAATGCCACCCATGCCCTGG
GGTGCCCCAGGGGACGTCCCCAGCTCCCGTGCCTTATGGTTACTTTGGAGGCGGGTACTA
CTCCTGCCGAGTGTCCCGGAGCTCGCTGAAACCCTGTGCCCAGGCAGCCACCCTGGCCGC
GTACCCCGCGGAGACTCCCACGGCCGGGGAAGAGTACCCCAGCCGCCCCACTGAGTTTGC
CTTCTATCCGGGATATCCGGGAACCTACCAGCCTATGGCCAGTTACCTGGACGTGTCTGT
GGTGCAGACTCTGGGTGCTCCTGGAGAACCGCGACATGACTCCCTGTTGCCTGTGGACAG
TTACCAGTCTTGGGCTCTCGCTGGTGGCTGGAACAGCCAGATGTGTTGCCA

BM794275

GCAGACTCTGGGTGCTCCTGGAGAACCGCGACGTGACTCCCTGTTGCCTGTGGACAGTTA
CCACTCTTGGGCTCTCGCTGGTGGCTGGAACAGCCAGATGTGTTGCCAGGGAGAACAGAA
CCCACCAGGTCCCTTTTGGAAGGCAGCATTTGCAGACTCCAGCGGGCAGCACCCTCCTGA
CGCCTGCGCCTTTCGTCGCGGCCGCAAGAAACGCATTCCGTACAGCAAGGGGCAGTTGCG
GGAGCTGGAGCGGGAGTATGCGGCTAACAAGTTCATCACCAAGGACAAGAGGCGCAAGAT
CTCGGCAGCCACCAGCCTCTCGGAGCGCCAGATTACCATCTGGTTTCAGAACCGCCGGGT
CAAAGAGAAGAAGGTTCTCGCCAAGGTGAAGAACAGCGCTACCCCTTAAGAGATCTCCTT
GCCTGGGTGGGAGGATCTAAAGTGGGGGTGTCCTGGGGAGACCAGGAACCTGCCAAGCCC
AGGCTGGGGCCAAGGACT

BQ363211

ACGCTGCACTGCGTTTCAAAGAGAAGAAGGTTCTCGCCAAGGTGAAGAACAGCGCTACCC
CTTAAGAGATCTCCTTGCTTGGGTGGGAGGAGCGAAAGTGGGGGTGTCCTGGGGAGACCA
GGAACCTGCCATCACCAGGCTGGGCCCAAGGACTCTGCTGAGAGGCCCCTAGAGACAACA
CCCTTCCCAGGCCATTGCTTGCTGGACTGTGCCTCAGGAGCGGCCTGGGTACC

BM932052

GAGTTTTCCAATTTCCAAAGAAAAATTTAGGTTTCCTGCAGCCGTGACATATGTGTGTGC
ACTGGGATGGGTTAATGTGTGTGTGTGTGTGTGTATGCGCATGTATTGGGAGTGGGGGCA
GAAACGTGTTTCCAGAATTTGCCTGTAGAATCTAAAAGAGTGGCCAAGAGTCTGGAAATG

CATGAAGACTGGACGTATGTGATGGTGGGCAAAGGCCTGACTGTGTGTGGTGTGTGGGTA
TGTTTGCAGATTCGCGGGTGTGAGAGCAGTGATGGGTGAGGGTGGCCTTCAGGAGCCAAG
GCTGATCGGTGGTGAGAGAACAAGCCGGAAGCCAGGGTGCTGTCCTGGTATGCTTTGGAG
GAACAGGATTGCACGTGCGCCTGTAGGGTGACCTGTGTGCACCTGTGAGATGACTTAGCT
TGGGGCTTGCAAGGCCTGGGTCTGCATGGGTGGGTATCTGACCATGCCTTTTCCTCCCTC
CCTTTCACGCCGCGCAGACTCCAGCGGGCAGCACCCTCCTGACGCCTGCGCCTTTCGTC

AA357646.1

CCGGCCTGGGTGGGGAGAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCGCACCTCAT
GAGCCGACCCTCGGCTCCATGGAGCCCGGCAATTATGCCACCTTGGATGGAGCCAAGGAT
ATCGAAGGCTTGCTGGGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCCCTCTGACC
AGCCACCCAGCGGCGCCTACGCTTGATGCCTGCTTGTCAACTATGCCCCCTTGGATCTGC

AW609525

ACCGCGGGTCAAATTTATTCATAATTAGCTCAATCATGAAAGCGGTTCTAAAGTGCTCTA
CAGAGCTCTAGATAGAAAATATGAGGCTAACGATCATGGCAGCTAGTACTGGTTATCGTG
ATTATGGCCACTGTCAGGATGAATGATAATGACTGGGCCAGGTCCTTTGGAAACCCTGGT
GGAGTGGGCTGTCACATGGGGTCCCGTCTCCCTGCACATACTGGGTACCCAGGCCGCTCC
TGAGGAACAGTCCAGCAGCCAGTGGCCTGGGAAGGGTGTGGTCTCTAGGGGCCTCTCAGC
AGAGTCCTTGGCCCCAGCCTGGGCTTGGCAGGTCCCTGGTCTCCCCAGGACACCCCCACT
TTCGCTCCTCCCACCCAGGCAAGGAGATCTCTTAAGGGGTAGCGCTGTTCTTCACCTTGG
CGAGAACCTTCTTCTCTTTGAACCGGCGGTGCGGCGTGGGGTACCGAGC

CB126919

ATTTTTCTGTCGTGTGAAAATGAAGCCAGCAGGCTGCCCCTAGTCAGTCCTTCCTTCCAG
AGAAAAAGAGATTTGAGAAAGTGCCTGGGTAATTCACCATTAATTTCCTCCCCCAAACTC
TCTGAGTCTTCCCTTAATATTTCTGGTGGTTCTGACCAAAGCAAGTCATGGTTTGTTGAG
CATTTGGGATCCCAGTGAAGTAGATGTTTGTAGCCTTGCATACTTAGCCCTTCCCAGGCA
CAAACGGAGTGGCAGAGTGGTGCCAACCCTGTTTTCCCAGTCCACGTAGACAGATTCACA
GTGCGGAATTCTGGAAGCTGGAGACAGACGGGCTCTTTGCAGAGCCGGGACTCTGAGAGG
GACATGAAGGCCTCTGCCTCTGTGTTCATTCTCTGATGTCCTGTACCTGGGCTCAGTGCC
CGGTGGGACTCATCTCCTGGCTGCGCAGCAAAGCCAGCGGGTTCGTGCTGGT

AW609336

CCAACGAGAAGAAGGTTCTCGCAAGGTGAAGAACAGCGCTACCCCTTAAGAGATCTCCTT
GCGTGGGTGGGAGGAGCGAAAGTGGGGGTGTCCTGGGGAGACCAGGAACCTGCCAGCCCA
GGCTGAGGCCAAGGACTCTGCTGAGAGGCCCCTAGAGACAACACCCTTCCCAGGCCACTG
GATGCTGAACTGTCCCTCAGGAGCGGCCTGGGTACCCAGTATGTGCAGGGAGACGGAACC
CCATGTGACAGCCCACTCCACCAGGGTTCCCAAAGAACCTGGCCCCAGTCATAATCATTC
ATCCTGACAGTGGCAATAATCACGATAACCAGTACTAGCTGCCATGATCGTAAGCCTCAT

ATTTGCTATCTAGAGCTCTGTAGAGCACTTTAGAAACCGCTTTCATGAATTGAGCTAATT
ATGACTCAATTTGAACCGGCGTCCGGCGTG

AW609244

ACGCGCACCGCGGTCAAGAGAAGAAGGTTCTCGCAAGGTGAAGAACAGCGCTACCCCTTA
AGAGATCTCCTTGCGTGGGTGGGAGGAGCGAAAGTGGGGGTGTCCTGGGGAGACCAGGAA
CCTGCCAAGCCCAGGCTGTGGCCAAGGACTCTGCTGAGAGGCCCCTATGAGACAACACCC
TTCCCAGGCCACTGGCTGCTGGGACTGTTCCTCAGGAGCGGCCTGGGTACCCGAGTAATG
TGCAGGGGAGACGGAACCCCATGTGACAGCCCACTCCACCAGGGTTCCCAAAAGAACCCT
GGCCCAGTCATAATCATTCATCCTGACAGTGGCAATAATCACGATAACCAGTACTAGCTG
CCATGATCGTAAGCCTCATATTTGCTATCTAGAGCTCTGTAGAGCCCTTTAGAAACCGCT
TTCATGAATGGAGCTAAATTATGAATACATTTGAACCGGCGATCCGACGTGA

BF855145

CTAGAGGATCCCGGAAGCAACTGCAACAGGTTCCCAAAGAACCGGGCCAGTCATAATCAT
TCATCCTGACAGGGCAATAATCACGATAACCAGTACTAGCTGCCATGATCGTTAGCCTCA
TATTTTCTATCTAGAGCTCTGTAGAGCACTTTAGAAACCGCTTTCATGAATGGAGCTAAT
TATGAATAAATTTGGAAGGCGATCCCTTGGCAGGGAAGCTTTCTCTCAGACCCCCTTCCA
TTACACCTCTCACCCTGGTAACAGCAGGAAGACTGAGGAGAGGGGAACGGGCAGATTCGT
GGTGTTGCAGTGTGCTTCCG

AU126914

GAGCGAATGCAGGCGACTTGCGAGCTGGGAGCGATTTAAAACGCTTTGGATTCCCCCGGC
CTGGGTGGGGAGAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCGCACCTCATGAGCC
GACCCTCGGCTCCATGGAGCCCGGCAATTATGCCACCTTGGATGGAGCCAAGGATATCGA
AGACTTGCTGGGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCCCTCTGACCAGCCA
CCCAGCGGCGCCTACGCTGATGCCTGCTGTCAACTATGCCCCCTTGGATCTGCCAGGCTC
GGCGGAGCCGCCAAAGCAATGCCACCCATGCCCTGGGGTGCCCCAGGGGACGTCCCCAGC
TCCCGTGCCTTATGGTTACTTTGGAGGCGGGTNCTACTCCTGCCGAGTGTCCCGGAGCTC
GCTGAAACCCTGTGCCCANNCANCCACCCTGGCCGCGTN

CB126449

CTCTGCCTCTGTGTTCATTCTCTGATGTCCTGTACCTGTGCTCAGTGCCCGGTGGGACTC
ATCTCCTGGCTGCGCAGCAAAGCCAGCGGGTTCGTGCTGGTCCTTCCTGCACCTTCGGCT
GGGGGTGGGGGGCCTGCCGGCGCATTCTCCACGATT

AW582404

ACGCTGCACCGCCGGTCCAAGAGAAGAAGGTTCTCGCCAAGGTGAAGAACAGCGCTACCC
CTTTAAGAGATCTCCTTGCTGGGGTGGGAGGAGCGAAAGTGGGGGTGTCTGGGGAGACCA

GGAACCTGCCAGCCCCAGGCTGGGCCCAAGGACTCTGCTGAGAGGCCCCTAGAGACAACA
CCCTTCCCAGGCCACTGTCTGCTGGACTGTTCCTCAGGAGCGGCCTGGGTACNCAGTATG
TGCAGGGAGACGGAACCCCATGTGACAGCCCACTCCACCAGGGTTCCCAAAGAACCTGGC
CCAGTCATAATCATTCATCCTGACAGTGGCAATAATCACGATAACCAGTACTAGCTGCCA
TGATCGTTAGCCTCATATTTTCTATCTAGAGCTCTGTAGAGCACTTTAGAAACCGCTTTC
ATGAATTGAGCTACTTATGAATCACTTTGAACCGGCGGTGCGGCGTG

BX641644

GGGGGAGAGCGAGCTGGGTGCCCCCTAGATTCCCCGCCCCCGCACCTCATGAGCCGACCC
TCGGCTCCATGGAGCCCGGCAATTATGCCACCTTGGATGGAGCCAAGGATATCGAAGGCT
TGCTGGGAGCGGGAGGGGGGCGGAATCTGGTCGCCCACTCCCCTCTGACCAGCCACCCAG
CGGCGCCTACGCTGACGCCTGCTGTCAACTATGCCCCCTTGGATCTGCCAGGCTCGGCGG
AGCCGCCAAAGCAATGCCACCCATGCCCTGGGGTGCCCCAGGGGACGTCCCCAGCTCCCG
TGCCTTATGGTTACTTTGGAGGCGGGTACTACTCCTGCCGAGTGTCCCGGAGCTCGCTGA
AACCCTGTGCCCAGGCAGCCACCCTGGCCGCGTACCCCGCGGAGACTCCCACGGCCGGGG
AAGAGTACCCCAGCCGCCCCACTGAGTTTGCCTTCTATCCGGGATATCCGGGAACCTACC
AGCCTATGGCCAGTTACCTGGACGTGTCTGTGGTGCAGACTCTGGGTGCTCCTGGAGAAC
CGCGACATGACTCCCTGTTGCCTGTGGACAGTTACCAGTCTTGGGCTCTCGCTNGTGGCT
GGAACAGCCAGATGTGTTGCCAGGGAGAACAGAACCCACCAGGTCCCTTTTGGAAGGCAG
CATTTG

Sequences from Table 4 not disclosed above

AW006861 (IMAGE Clone ID: :2497262)

GCTGAGTTCTGAAGCTTCTGAGTTCTGCAGCCTCACCTCTGAGAAAACCTCTTTTCCACC
AATACCATGAAGCTCTGCGTGACTGTCCTGTCTCTCCTCATGCTAGTAGCTGCCTTCTGC
TCTCTAGCGCTCTCAGCACCAATGGGCTCAGACCCTCCCACCGCCTGCTGCTTTTCTTAC
ACCGCGAGGAAGCTTCCTCGCAACTTTGTGGTAGATTACTATGAGACCAGCAGCCTCTGC
TCCCAGCCAGCTGTGGTATTCCAAACCAAAAGAAGCAAGCAAGTCTGTGCTGATCCCAGT
GAATCCTGGGTCCAGGAGTACGTGTATGACCTGGAACTGAACTGAGCTGCTCAGAGACAG
GAAGTCTTCAGGGAAGGTCACCTGAGCCCGGATGCTTCTCCATGAGACACATCTCCTCCA
TACTCAGGACTCCTCTCCGCAGTTCCTGTCCCTTCTCTTAATTTAATCTTTTTTATGTGC
CGTGTTATTGTATTAGGTGTCATTTCCATTATTTATATTAGTTTAGCCAAAGGATAAGTG
TCCCCTATGGGGATGGTCCACTGTCACTGTTTCTCTGCTGTTGCAAATACATGGATAACA
CATTTGATTCTGTGTGTTTTCATAATAAAACTTTAAAATAAAATGCAAAAAAAAAAAAAA
AAAA

X59770

GCCACGTGCTGCTGGGTCTCAGTCCTCCACTTCCCGTGTCCTCTGGAAGTTGTCAGGAGC
AATGTTGCGCTTGTACGTGTTGGTAATGGGAGTTTCTGCCTTCACCCTTCAGCCTGCGGC

ACACACAGGGGCTGCCAGAAGCTGCCGGTTTCGTGGGAGGCATTACAAGCGGGAGTTCAG
GCTGGAAGGGGAGCCTGTAGCCCTGAGGTGCCCCCAGGTGCCCTACTGGTTGTGGGCCTC
TGTCAGCCCCCGCATCAACCTGACATGGCATAAAAATGACTCTGCTAGGACGGTCCCAGG
AGAAGAAGAGACACGGATGTGGGCCCAGGACGGTGCTCTGTGGCTTCTGCCAGCCTTGCA
GGAGGACTCTGGCACCTACGTCTGCACTACTAGAAATGCTTCTTACTGTGACAAAATGTC
CATTGAGCTCAGAGTTTTTGAGAATACAGATGCTTTCCTGCCGTTCATCTCATACCCGCA
AATTTTAACCTTGTCAACCTCTGGGGTATTAGTATGCCCTGACCTGAGTGAATTCACCCG
TGACAAAACTGACGTGAAGATTCAATGGTACAAGGATTCTCTTCTTTTGGATAAAGACAA
TGAGAAATTTCTAAGTGTGAGGGGGACCACTCACTTACTCGTACACGATGTGGCCCTGGA
AGATGCTGGCTATTACCGCTGTGTCCTGACATTTGCCCATGAAGGCCAGCAATACAACAT
CACTAGGAGTATTGAGCTACGCATCAAGAAAAAAAAAGAAGAGACCATTCCTGTGATCAT
TTCCCCCCTCAAGACCATATCAGCTTCTCTGGGGTCAAGACTGACAATCCCGTGTAAGGT
GTTTCTGGGAACCGGCACACCCTTAACCACCATGCTGTGGTGGACGGCCAATGACACCCA
CATAGAGAGCGCCTACCCGGGAGGCCGCGTGACCGAGGGGCCACGCCAGGAATATTCAGA
AAATAATGAGAACTACATTGAAGTGCCATTGATTTTTGATCCTGTCACAAGAGAGGATTT
GCACATGGATTTTAAATGTGTTGTCCATAATACCCTGAGTTTTCAGACACTACGCACCAC
AGTCAAGGAAGCCTCCTCCACGTTCTCCTGGGGCATTGTGCTGGCCCCACTTTCACTGGC
CTTCTTGGTTTTGGGGGGAATATGGATGCACAGACGGTGCAAACACAGAACTGGAAAAGC
AGATGGTCTGACTGTGCTATGGCCTCATCATCAAGACTTTCAATCCTATCCCAAGTGAAA
TAAATGGAATGAAATAATTCAAACACAAAAAAAAAAAAAAAAAAAAA

AB000520

GGATCCAAGCTATTGTCCTGCCCATGGCTTCCCATCTCAGGACGCTCTCTGGCCGCTATC
ATCCCAGCAGTGGAGTTCAGCCCACTACTCTGAACCAGCCGCAGGTGGCTGCTATGGGAC
TGAAGCCATGAATGGTGCCGGCCCTGGCCCCGCCGCAGCCGCCCCGGTCCCAGTCCCGGT
CCCGGTCCCGGACTGGCGGCAGTTCTGCGAGCTGCATGCGCAGGCGGCCGCCGTGGACTT
TGCGCACAAGTTCTGCCGTTTCCTGCGGGACAACCCAGCTTACGACACGCCCGACGCCGG
CGCCTCCTTCTCCCGCCACTTCGCCGCCAACTTCCTGGACGTCTTCGGCGAGGAGGTGCG
CCGCGTGCTGGTGGCTGGGCCGACGACTCGGGGCGCGGCCGTGAGCGCAGAGGCCATGGA
GCCGGAGCTCGCGGACACCTCTGCACTCAAGGCGGCGTCCTACGGCCACTCGCGGAGCTC
GGAGGACGTGTCCACGCACGCGGCCACCAAGGCCCGCGTTCGCAAGGGCTTCTCGCTGCG
CAACATGAGCCTGTGCGTGGTGGACGGCGTGCGCGACATGTGGCACCGGCGCGCCTCGCC
CGAGCCCGACGCGGCAGCTGCCCCGCGCACCGCCGAGCCCCGCGACAAGTGGACGCGGCG
CCTGAGGCTGTCGCGGACGCTGGCTGCCAAGGTGGAGCTGGTGGACATTCAACGCGAGGG
GGCGCTGCGCTTCATGGTGGCCGACGACGCGGCCGCGGGCTCCGGGGGCTCGGCTCAGTG
GCAGAAGTGCCGCCTGCTCCTGCGCAGGGCTGTGGCCGAGGAACGCTTCCGCCTGGAGTT
CTTCGTGCCGCCCAAAGCCTCCAGGCCCAAGGTCAGCATCCCACTGTCAGCCATCATTGA
GGTCCGCACCACCATGCCCCTGGAAATGCCAGAGAAGGATAACACATTCGTCCTCAAGGT
AGAGAATGGAGCCGAATACATCTTGGAGACCATCGACTCTCTGCAGAAGCACTCGTGGGT
AGCTGACATCCAGGGCTGCGTGGACCCCGGTGACAGTGAGGAAGACACCGAGCTCTCCTG
TACCCGAGGAGGCTGTCTGGCCAGCCGCGTGGCCTCCTGCAGCTGTGAGCTCCTGACTGA
TGCAGTCGACCTGCCCCGCCCCCCAGAGACGACAGCCGTGGGTGCAGTGGTGACAGCCCC
CCACAGCCGAGGTCGAGATGCCGTCAGAGAATCCCTGATCCACGTCCCGCTAGAGACCTT

TCTGCAGACCCTGGAATCCCCGGGCGGCAGCGGCAGTGACAGCAATAACACAGGGGAACA
GGGTGCAGAGACGGATCCCGAGGCTGAACCCGAGCTGGAGCTATCCGACTACCCATGGTT
CCACGGGACACTGTCCCGGGTCAAGGCTGCTCAACTGGTTCTGGCAGGGGGGCCCCGGAA
CCACGGCCTCTTCGTGATCCGCCAAAGTGAGACTCGGCCTGGGGAGTACGTGCTGACCTT
CAACTTCCAGGGCAAGGCCAAGCACCTGCGCCTGTCCCTGAACGGCCACGGCCAGTGTCA
CGTACAGCATCTGTGGTTCCAGTCTGTGCTTGACATGCTCCGCCACTTCCACACACACCC
CATCCCACTGGAGTCAGGGGGCTCGGCCGACATCACCCTTCGCAGCTATGTGCGGGCCCA
GGACCCCCCACCAGAGCCGGGCCCCACGCCCCCTGCCGCGCCCGCGTCCCCGGCCTGCTG
GAGCGACTCGCCCGGCCAGCACTACTTCTCCAGCCTCGCCGCGGCCGCCTGCCCGCCTGC
CTCGCCCTCCGACGCCGCCGGCGCCTCCTCGTCTTCCGCCTCGTCGTCCTCTGCCGCGTC
GGGGCCCGCCCCCCCGCGCCCCGTCGAGGGCCAGCTCAGCGCGCGGAGCCGCAGCAACAG
CGCCGAGCGCCTGCTGGAGGCCGTGGCCGCCACCGCCGCCGAGGAGCCCCCGGAGGCCGC
GCCCGGCCGCGCGCGCGCCGTGGAGAACCAGTACTCCTTCTACTAGCCCGCGGCGCCGCC
CGGGTGGGACACGCCAAGCTCTTCAGTGAAGACACGATGTTATTAAAAGCCTGTTTTAGG
GACTGCAAAA

AI820604 (IMAGE Clone Id: 1605108

GATTCCAGCACGGGCTTCGCAGACTGCAGGACACAGAGGCACGCGTGCACATCATGTCTT
CTAAGGAATTTGAACACTGTTGAGAAGACTGTGTACAAGAGAGATGTGCCATGTCAGCCT
TGCAAGGGACAGCGTGAAAACTACCCATCTCCGGTCACCAAGTTGCAGGAGGCCAGGAGC
CAGGAGGGGAAACCGCTCAGTTTGCAAAACGTCGCTTCCACAAGCCTGATGGCTGAAACT
GCTCACTGTACCCTGAAACCAGCTTTACCTACAGCTTCTGAGATAAACTGCTGCAACTCT
GGGACCCACGATGCCTATCACAGTGGCTCATCAATGGAACCTGCCGGCTCCCAACCCTTC
CTAGGGCCCATGAACTCTCTGAAAAGAGGAACAGAAATATTTCTCCTTTTTGTAAAATCT
TTAACCTTCCCTTTGTTCTTCATGTACACGCTGAACTGCAATTCTTCTTCCCAAATAAAA
CATTAAATTTAAAAAA

AI087057 (IMAGE Clone ID:1671188)

GGCCCCGGAGGGAGAGTAACCCGGCCCATCCATCCGTCGCCCGGTTCTTGGGGAACTACT
TTCAGGGGCTTCTTGCCGTCCCCTCATCAGCTCTGTGCGAACCCTCTGTCGGCAGCCATT
GAGGAGACCCTGCCCCCTGGACCCTGACCACATATAGATTGAGGCCGAGGAGTGGCTGCC
CTGTCCCTTTTATGACAGCCCGCAGAAGCCCCGGGGTGAGGCATGGAGGAGGCAGGCGAC
AGCTGACAGGGACCCTGTTGGCCTCCAGCATGTCCAGCCAGCCGGGCAGGATTTCTCTGC
TTCTGGCTGGCAGCCAGGAACTGAGTATGACAATGTTGTACTAAAGAAAGGCCCAAAGTG
ACAGAGGCAGCAGAGGGATGGTCCACCGCCCCTTGGCTTCTGCTGGTGACTCCTCCTGGC
CACTGCATCAGAAGAACCTCCTCTGCCCCTTCTGGAGCCCGAGGCCTGGCCTGTCTTCGT
TGGGGCTGATAAATTGCCTCTCCCAGGGCCTGCTGGGTGAGTCACCATCCCAAAGCAGGA
AGGGTGCCCTGGAGAGAACCACCCTCCTCCTACTCTTTTTCCACTTCCTCCTCTTTCTTT
CCCCAGCTGAGGAGGAACCTGGGGCATTTAGGGCAGAGGACAAAAGGATGTCAGCAATTG
CTTGGGCTGCTTGGCTATGCAAGCCTCCTGCCTGCTGATGGCCACTTCAGGGACAGCCTG
GGCCCAGGCACCCAGGGGGATGGCGGCAGCTTCCTGCACCTTTCAGATTTCTTGGTGGCA
TTAAAGCATTTTCAGAACAAAAAAAAAAAAAAAAAAAAAAAA

AJ272267

GGCGGGCCTGGACGGCCGCGTGCTGTACTGGCCACGCGGCCGCGTCTGGGGTGGCTCCTC
ATCCCTCAATGCCATGGTCTACGTCCGTGGGCACGCCGAGGACTACGAGCGCTGGCAGCG
CCAGGGCGCCCGCGGCTGGGACTACGCGCACTGCCTGCCCTACTTCCGCAAGGCGCAGGG
CCACGAGCTGGGCGCCAGCCGGTACCGGGGCGCCGATGGCCCGCTGCGGGTGTCCCGGGG
CAAGACCAACCACCCGCTGCACTGCGCATTCCTGGAGGCCACGCAGCAGGCCGGCTACCC
GCTCACCGAGGACATGAATGGCTTCCAGCAGGAGGGCTTCGGCTGGATGGACATGACCAT
CCATGAAGGCAAACGGTGGAGCGCGGCCTGTGCCTACCTGCACCCAGCACTGAGCCGCAC
CAACCTCAAGGCCGAGGCCGAGACGCTTGTGAGCAGGGTGCTATTTGAGGGCACCCGTGC
AGTGGGCGTGGAGTATGTTAAGAATGGCCAGAGCCACAGGGCTTATGCCAGCAAGGAGGT
GATTCTGAGTGGAGGTGCCATCAACTCTCCACAGCTGCTCATGCTCTCTGGCATCGGGAA
TGCTGATGACCTCAAGAAACTGGGCATCCCTGTGGTGTGCCACCTACCTGGGGTTGGCCA
GAACCTGCAAGACCACCTGGAGATCTACATTCAGCAGGCATGCACCCGCCCTATCACCCT
CCATTCAGCACAGAAGCCCCTGCGGAAGGTCTGCATTGGTCTGGAGTGGCTCTGGAAATT
CACAGGGGAGGGAGCCACTGCCCATCTGGAAACAGGTGGGTTCATCCGCAGCCAGCCTGG
GGTCCCCCACCCGGACATCCAGTTCCATTTCCTGCCATCCCAAGTGATTGACCACGGGCG
GGTCCCCACCCAGCAGGAGGCTTACCAGGTACATGTGGGGCCCATGCGGGGCACGAGTGT
GGGCTGGCTCAAACTGAGAAGTGCCAATCCCCAAGACCACCCTGTGATCCAGCCCAACTA
CTTGTCAACAGAAACTGATATTGAGGATTTCCGTCTGTGTGTGAAGCTCACCAGAGAAAT
TTTTGCACAGGAAGCCCTGGCTCCGTTCCGAGGGAAAGAGCTCCAGCCAGGAAGCCACAT
TCAGTCAGATAAAGAGATAGATGCCTTTGTGCGGGCAAAAGCCGACAGCGCCTACCACCC
CTCGTGCACCTGTAAGATGGGCCAGCCCTCCGATCCCACTGCCGTGGTGGATCCGCAGAC
AAGGGTCCTCGGGGTGGAAAACCTCAGGGTCGTCGATGCCTCCATCATGCCTAGCATGGT
CAGCGGCAACCTGAACGCCCCCACAATCATGATCGCAGAGAAGGCAGCTGACATTATCAA
GGGGCAGCCTGCACTCTGGGACAAAGATGTCCCTGTCTACAAGCCCAGGACGCTGGCCAC
CCAGCGCTAAGACAGTTGCTGCTGGAGGATGACCAGGGAAGCCCCCTGATAAGCCAAGAG
GGCCAGCACAGCCCTTGCTCCCAGGCTCCTGCCTGAAACTATCTAGCACACTAGGACCCA
GGTGGTACCCTACTCAGTGGCTGAGAATTGGATAAAGTCTTKGGGAAATGAGACAAGTAC
TGGGCAGTGAATCCAGCTCCTTTTCCCCAGCCTTTCCCTGTGGGCCATTTGGGGAAGGCC
AGCATTYCAGCCTGAGATGTTCCTCCCTGCCTCCTGGGGGGGCARAAGGGVTAGGWTGGT
TAACTCCTGCCGCATCCTTCCCTGCCTCCTGGAGGGACAGAAGGGGAGGATGGTTAACTC
CTGCCGCATCCTTTTTCTTGTGTTCACGTGGCATTCTCTAACCCAGGGCAGTGGTTCCTT
CCCAGGCCATGCACAGAGGCTGGGTGCCTGCCAGACCCACGGAGGGTTCGCGAAGGAAGG
GGCATCCTCCTTCTTGAGCTGCAAGCTTTAGCTGAGGCAGTAAGTCACACAGTAGTTAGT
TCAGCCTGGGCTGGCACATAAGTCCCCAGTGTCCCTGTTGAGAGGGGAAAGTTGCCTGCT
GGTTGAAAAACTGGCTTTTCCTTTCTCGCTGCCTAATTTCACTCTCAGAGTGAGGCAGGT
AACTGGGGCTCCACTGGGTCACTCTGAGAGGGTTGTGGCTCTGGTTCTTATTAAACCAGG
GCCAGGTGCAGGGCTCACACCTGTAATCCCAGCACTTTGGGAAGGTCACTTGAGCTCAGG
AGTTCAAGACCAGCCTGGGCAACATAGTGAGACCTTGTCTCTGGAAAACAATTAGCTGGG
CATGGTGGTACACACCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCGGGAGGATGGCTTT
AGCCCAGGAGGTTGAGGCTCCTGTGAACCCTGATGGCACCACTGCACTCCAGCCTGGGTG
ACAGGGTGAGACCCTGTCTCAAAAAAAAA

N30081 (IMAGE Clone ID: 258695)

CCGCCGTTGNCAAAGGGCCCAGAATATGGGCCATGGACNATCTCCATGCCTGGGGAAATT
CCCTCGGGTCTTTTGGNTAACCNCCTTATAGAAAGGTAATGNCATGGAGTCTCTACAGGG
NGCACAAGGTGGACTAATTGATACGAAGAGCCCTGTAAATATGTGGGCAGCGGCAGATTT
TGACCATTTGGACCGAACTGTATTTGACACAGCGCAATATCTGGAACTGGTTGGTCAAAA
ACCTGCTTGTCTTGTTAAATTTCCTCTGTCCAAGGACATGGAATCTCTCTCTAATTTTAC
TTCAAATTTCCCTTTCCTTCATTTCTCTAAAAACGTTAAATAAGAAAGAAGATTGTAAAG
CCAGCATTTGAAGCCTAAGTATTGAAAGTCTTTGACAATTTCTGAAATCAGACTTGACAT
CTTTCCCCCGCCTTGCAAATTTCTTGAAGAAATAAGAAGCTACATGTAAGCATCATCATG
TTTATTAAATTACAATGAGAACTCTCACTCAATCTTGACCAGAGCAGACTCTTAACTTGG
AAGCAGAGTCCCTCTAAAGGTAACTCTTGTGGTCACTCAATATTGTATTGGCATTTGCAT
ATTAAATAGACATTTCAGTAGCATTT

AI700363 (IMAGE Clone ID: 2327403)

TGGCCCGCGGTCGCGGTGGGATCCTAGCCCTGTCTCCTCTCCTGGGAAGGAGTGAGGGTG
GGACGTGACTTAGACACCTACAAATCTATTTACCAAAGAGGAGCCCGGGACTGAGGGAAA
AGGCCAAAGAGTGTGAGTGCATGCGGACTGGGGGTTCAGGGGAAGAGGACGAGGAGGAGG
AAGATGAGGTCGATTTCCTGATTTAAAAAATCGTCCAAGCCCCGTGGTCCAGCTTAAGGT
CCTCGGTTACATGCGCCGCTCAGAGCAGGTCACTTTCTGCCTTCCACGTCCTCCTTCAAG
GAAGCCCCATGTGGGTAGCTTTCAATATCGCAGGTTCTTACTCCTCTGCCTCTATAAGCT
CAAACCCACCAACGATCGGGCAAGTAAACCCCCTCCCTCGCCGACTTCGGAACTGGCGAG
AGTTCAGCGCAGATGGGCCTGTGGGGAGGGGGCAAGATAGATGAGGGGGAGCGGCATGGT
GCGGGGTGACCCCTTGGAGAGAGGAAAAAGGCCACAAGAGGGGCTGCCACCGCCACTAAC
GGAGATGGCCCTGGTAGAGACCTTTGGGGGTCTGGAACCTCTGGACTCCCCATGCTCTAA
CTCCCACACTCTGCTATCAGAAACTTAAACTTGAGGATTTTCTCTGTTTTTCACTCGCAA
TAAATTCAGAGCAAACAAAAAAAAAAAAAAA

AL117406

CAATAGGCCGGCTTTTGAACTGCTTCGCAGGGGACTTGGAACAGCTGGACCAGCTCTTGC
CCATCTTTTCAGAGCAGTTCCTGGTCCTGTCCTTAATGGTGATCGCCGTCCTGTTGATTG
TCAGTGTGCTGTCTCCATATATCCTGTTAATGGGAGCCATAATCATGGTTATTTGCTTCA
TTTATTATATGATGTTCAAGAAGGCCATCGGTGTGTTCAAGAGACTGGAGAACTATAGCC
GGTCTCCTTTATTCTCCCACATCCTCAATTCTCTGCAAGGCCTGAGCTCCATCCATGTCT
ATGGAAAAACTGAAGACTTCATCAGCCAGTTTAAGAGGCTGACTGATGCGCAGAATAACT
ACCTGCTGTTGTTTCTATCTTCCACACGATGGATGGCATTGAGGCTGGAGATCATGACCA
ACCTTGTGACCTTGGCTGTTGCCCTGTTCGTGGCTTTTGGCATTTCCTCCACCCCCTACT
CCTTTAAAGTCATGGCTGTCAACATCGTGCTGCAGCTGGCGTCCAGCTTCCAGGCCACTG
CCCGGATTGGCTTGGAGACAGAGGCACAGTTCACGGCTGTAGAGAGGATACTGCAGTACA
TGAAGATGTGTGTCTCGGAAGCTCCTTTACACATGGAAGGCACAAGTTGTCCCCAGGGGT
GGCCACAGCATGGGGAAATCATATTTCAGGATTATCACATGAAATACAGAGACAACACAC

CCACCGTGCTTCACGGCATCAACCTGACCATCCGCGGCCACGAAGTGGTGGGCATCGTGG
GAAGGACGGGCTCTGTAGGTTTTTACTGAGCACCTACTATGTGCCTGGGAACCGAAAGGG
AAGTCCTCCTTGGGCATGGCTCTCTTCCGCCTGGTGGAGCCCATGGCAGGCCGGATTCTC
ATTGACGGCGTGGACATTTGCAGCATCGGCCTGGAGGACTTGCGGTCCAAGCTCTCAGTG
ATCCCTCAAGATCCAGTGCTGCTCTCAGGAACCATCAGATTCAACCTAGATCCCTTTGAC
CGTCACACTGACCAGCAGATCTGGGATGCCTTGGAGAGGACATTCCTGACCAAGGCCATC
TCAAAGTTCCCCAAAAAGCTGCATACAGATGTGGTGGAAAACGGTGGAAACTTCTCTGTG
GGGGAGAGGCAGCTGCTCTGCATTGCCAGGGCTGTGCTTCGCAACTCCAAGATCATCCTT
ATCGATGAAGCCACAGCCTCCATTGACATGGAGACAGACACCCTGATCCAGCGCACAATC
CGTGAAGCCTTCCAGGGCTGCACCGTGCTCGTCATTGCCCACCGTGTCACCACTGTGCTG
AACTGTGACCACATCCTGGTTATGGGCAATGGGAAGGTGGTAGAATTTGATCGGCCGGAG
GTACTGCGGAAGAAGCCTGGGTCATTGTTCGCAGCCCTCATGGCCACAGCCACTTCTTCA
CTGAGATAAGGAGATGTGGAGACTTCATGGAGGCTGGCAGCTGAGCTCAGAGGTTCACAC
AGGTGCAGCTTCGAGGCCCACAGTCTGCGACCTTCTTGTTTGGAGATGAGAACTTCTCCT
GGAAGCAGGGGTAAATGTAGGGGGGGTGGGGATTGCTGGATGGAAACCCTGGAATAGGCT
ACTTGATGGCTCTCAAGACCTTAGAACCCCAGAACCATCTAAGACATGGGATTCAGTGAT
CATGTGGTTCTCCTTTTAACTTACATGCTGAATAATTTTATAATAAGGTAAAAGCTTATA
GTTTTCTGATCTGTGTTAGAAGTGTTGCAAATGCTGTACTGACTTTGTAAAATATAAAAC
TAAGGAAAACTCAAAAAAAAAAAA

M92432

CCCACAGGGGGACCGGCCCTGTGACCCCTCACCGGGGCCGTGGGCCCGAGCCCCGGACTT
CCCTAAGCCGGCAATGACCGCCTGCGCCCGCCGAGCGGGTGGGCTTCCGGACCCCGGGCT
CTGCGGTCCCGCGTGGTGGGCTCCGTCCCTGCCCCGCCTCCCCCGGGCCCTGCCCCGGCT
CCCGCTCCTGCTGCTCCTGCTTCTGCTGCAGCCCCCCGCCCTCTCCGCCGTGTTCACGGT
GGGGGTCCTGGGCCCCTGGGCTTGCGACCCCATCTTCTCTCGGGCTCGCCCGGACCTGGC
CGCCCGCCTGGCCGCCGCCCGCCTGAACCGCGACCCCGGCCTGGCAGGCGGTCCCCGCTT
CGAGGTAGCGCTGCTGCCCGAGCCTTGCCGGACGCCGGGCTCGCTGGGGGCCGTGTCCTC
CGCGCTGGCCCGCGTGTCGGGCCTCGTGGGTCCGGTGAACCCTGCGGCCTGCCGGCCAGC
CGAGCTGCTCGCCGAAGAAGCCGGGATCGCGCTGGTGCCCTGGGGCTGCCCCTGGACGCA
GGCGGAGGGCACCACGGCCCCTGCCGTGACCCCCGCCGCGGATGCCCTCTACGCCCTGCT
TCGCGCATTCGGCTGGGCGCGCGTGGCCCTGGTCACCGCCCCCCAGGACCTGTGGGTGGA
GGCGGGACGCTCACTGTCCACGGCACTCAGGGCCCGGGGGCTGCCTGTCGCCTCCGTGAC
TTCCATGGAGCCCTTGGACCTGTCTGGAGCCCGGGAGGCCCTGAGGAAGGTTCGGGACGG
GCCCAGGGTCACAGCAGTGATCATGGTGATGCACTCGGTGCTGCTGGGTGGCGAGGAGCA
GCGCTACCTCCTGGAGGCCGCAGAGGAGCTGGGCCTGACCGATGGCTCCCTGGTCTTCCT
GCCCTTCGACACGATCCACTACGCCTTGTCCCCAGGCCCGGAGGCCTTGGCCGCACTCGC
CAACAGCTCCCAGCTTCGCAGGGCCCACGATGCCGTGCTCACCCTCACGCGCCACTGTCC
CTCTGAAGGCAGCGTGCTGGACAGCCTGCGCAGGGCTCAAGAGCGCCGCGAGCTGCCCTC
TGACCTCAATCTGCAGCAGGTCTCCCCACTCTTTGGCACCATCTATGACGCGGTCTTCTT
GCTGGCAAGGGGCGTGGCAGAAGCGCGGGCTGCCGCAGGTGGCAGATGGGTGTCCGGAGC
AGCTGTGGCCCGCCACATCCGGGATGCGCAGGTCCCTGGCTTCTGCGGGGACCTAGGAGG
AGACGAGGAGCCCCCATTCGTGCTGCTAGACACGGACGCGGCGGGAGACCGGCTTTTTGC

CACATACATGCTGGATCCTGCCCGGGGCTCCTTCCTCTCCGCCGGTACCCGGATGCACTT
CCCGCGTGGGGGATCAGCACCCGGACCTGACCCCTCGTGCTGGTTCGATCCAAACAACAT
CTGCGGTGGAGGACTGGAGCCGGGCCTCGTCTTTCTTGGCTTCCTCCTGGTGGTTGGGAT
GGGGCTGGCTGGGGCCTTCCTGGCCCATTATGTGAGGCACCGGCTACTTCACATGCAAAT
GGTCTCCGGCCCCAACAAGATCATCCTGACCGTGGACGACATCACCTTTCTCCACCCACA
TGGGGGCACCTCTCGAAAGGTGGCCCAGGGGAGTCGATCAAGTCTGGGTGCCCGCAGCAT
GTCAGACATTCGCAGCGGCCCCAGCCAACACTTGGACAGCCCCAACATTGGTGTCTATGA
GGGAGACAGGGTTTGGCTGAAGAAATTCCCAGGGGATCAGCACATAGCTATCCGCCCAGC
AACCAAGACGGCCTTCTCCAAGCTCCAGGAGCTCCGGCATGAGAACGTGGCCCTCTACCT
GGGGCTTTTCCTGGCTCGGGGAGCAGAAGGCCCTGCGGCCCTCTGGGAGGGCAACCTGGC
TGTGGTCTCAGAGCACTGCACGCGGGGCTCTCTTCAGGACCTCCTCGCTCAGAGAGAAAT
AAAGCTGGACTGGATGTTCAAGTCCTCCCTCCTGCTGGACCTTATCAAGGGAATAAGGTA
TCTGCACCATCGAGGCGTGGCTCATGGGCGGCTGAAGTCACGGAACTGCATAGTGGATGG
CAGATTCGTACTCAAGATCACTGACCACGGCCACGGGAGACTGCTGGAAGCACAGAAGGT
GCTACCGGAGCCTCCCAGAGCGGAGGACCAGCTGTGGACAGCCCCGGAGCTGCTTAGGGA
CCCAGCCCTGGAGCGCCGGGGAACGCTGGCCGGCGACGTCTTTAGCTTGGCCATCATCAT
GCAAGAAGTAGTGTGCCGCAGTGCCCCTTATGCCATGCTGGAGCTCACTCCCGAGGAAGT
GGTGCAGAGGGTGCGGAGCCCCCCTCCACTGTGTCGGCCCTTGGTGTCCATGGACCAGGC
ACCTGTCGAGTGTATCCTCCTGATGAAGCAGTGCTGGGCAGAGCAGCCGGAACTTCGGCC
CTCCATGGACCACACCTTCGACCTGTTCAAGAACATCAACAAGGGCCGGAAGACGAACAT
CATTGACTCGATGCTTCGGATGCTGGAGCAGTACTCTAGTAACCTGGAGGATCTGATCCG
GGAGCGCACGGAGGAGCTGGAGCTGGAAAAGCAGAAGACAGACCGGCTGCTTACACAGAT
GCTGCCTCCGTCTGTGGCTGAGGCCTTGAAGACGGGGACACCAGTGGAGCCCGAGTACTT
TGAGCAAGTGACACTGTACTTTAGTGACATTGTGGGCTTCACCACCATCTCTGCCATGAG
TGAGCCCATTGAGGTTGTGGACCTGCTCAACGATCTCTACACACTCTTTGATGCCATCAT
TGGTTCCCACGATGTCTACAAGGTGGAGACAATAGGGGACGCCTATATGGTGGCCTCGGG
GCTGCCCCAGCGGAATGGGCAGCGACACGCGGCAGAGATCGCCAACATGTCACTGGACAT
CCTCAGTGCCGTGGGCACTTTCCGCATGCGCCATATGCCTGAGGTTCCCGTGCGCATCCG
CATAGGCCTGCACTCGGGTCCATGCGTGGCAGGCGTGGTGGGCCTCACCATGCCGCGGTA
CTGCCTGTTTGGGGACACGGTCAACACCGCCTCGCGCATGGAGTCCACCGGGCTGCCTTA
CCGCATCCACGTGAACTTGAGCACTGTGGGGATTCTCCGTGCTCTGGACTCGGGCTACCA
GGTGGAGCTGCGAGGCCGCACGGAGCTGAAGGGCAAGGGCGCCGAGGACACTTTCTGGCT
AGTGGGCAGACGCGGCTTCAACAAGCCCATCCCCAAACCGCCTGACCTGCAACCGGGGTC
CAGCAACCACGGCATCAGCCTGCAGGAGATCCCACCCGAGCGGCGACGGAAGCTGGAGAA
GGCGCGGCCGGGCCAGTTCTCTTGAGAAGTGAGGCCCGGCCCCGGACAGGGTCTGGGCCC
TGCTCCCTGTCCCATCTGCAGTGGACCCCAGGCACCCCCCTTTGAGGAGGTGGGGTGAAC
TGCTCCTTGGCAGGGATTTGTGACACTGCATTGCTGGGCTGTGTTCCTCGGGCTCTTCTG
GACCTTGCACCGTGGATACCAGGCCATGTGCCATGGTATTTGGGTCCTGGGAGGGTGGGT
GAAATAAAGGCATACTGTCTT

AL050227

CTTTCACAGAAAGAAAGTAACAGGCATAATTCCTGTTGATGAGGCTGGGATTGTTTTTAA
GAGGAGAGATAATAACTTCATATTTTTAAAGTGCCAGTAGCCTAATATGTGAAACAGATC

AGAATCTGTTGTGTAGTAAGTCTGCTTTGTTGAAGAATTTATTATGGGAGTAAAGATAAG
AAGGAAAGAGATCACCATCAGAAACAAGTCAGCCTTTTCATGCTTTTTTGAGCATTTTTG
GAGATGATTCCACTTCTCAAGTTATTATCATTTGTGCATCTCTTCAATGCTATTGTTAAA
TGCTTTAGAATTAGAATATTTTGATCCTTTAATTAAAGTAAGCCAAACGTCTAGGCAAAA
ACAGCCAATCATTAAACTTTAATAGTAATTCAAATATAGATTTCTCATACAGTTTTCCAT
GTCTGTAGAAATCAAAGTTGTAATGTTAAGCAGAGGGAAATGCGTGTGATTTACTAATAC
ACTTCAACGTTCTACTTTTGAAAGGATACTCATGTGGGTGGGGCAGAGAACATAGAAAAA
GATATGATGGAAAACCTGTCCATTTTCTACCTGTTAACCTTCATCATTTTGTGCAGGCCC
TGGAAGCAAAGAGAGGAAGGGACCGACTGCATTTATCTTTGAACACTTGAGCATCAGTAG
TACTACTGAGTGGCCAGGGGTCTTGTCTGTCAAAGCAAATGATAAGTTCACTCAGGCCAT
TATTGACTGCTGAACTCTCTTCCTTCCCAACTCTTCCTTGAAAGAGAAAAAAATACTTTG
CCTTCTTGCTCTCCTTATCAAATGTTTTTGTACAAATAGTGTAAGCCTGTTTAAGCAAAC
CAATTAAAATAGGCACTGATTATTTTGATCTGTTTGTAACAAATGAATGTAAGTACTATT
TACATGGTGTGCCTAGGAGGAGCTGAAATCATTGGCACTTTAATCCATATTGTAAAGATC
AGTATCAAAAGCATAGTGTTCTTCACCTCTCCTCCTCAGCATCCATCTCTATATACTTGA
TTAAATGGAAAAGTCTCTTTTATCACCTCTATGTAAAGTTTTATGGGTAGTTATCGTCAG
TGTATTTAAATATATCTTCTAGTATGTTTTAAAGGCTGGTCTTCAATACTGTGGAGACAA
AAAATAAAAGAGCGTATGAAAAGTACGTTAGACTTTTGCTGGCATTCAAGTCATGGCTAG
TCTGTGTATTTAATAAATGTGTGTTATTTATGTCGTGTTTGTCAATGGAAAATAAAGTTG
AATATTCTGAAAAAAAAAAAAAAA

AW613732 (IMAGE Clone ID: 2953502)

CCTANAAGTNCCATTTTGGCAAGGATAAACTCCCATGACAANCTCCCANTACTGCATGTG
AATGAATAAGAAACAAGAANTGACCACACCAAAGCCTCCCTGGCTGGTGTTACANGGGAT
CAGGTCCACAGTGGTGCAGATTCAACCACCACCCAGGGAGTGCTTGCAGACTCTGCATAG
ATGTTGCTGCATGCGTCCCATGTGCCTGTCAGAATGGCAGTGTTTAATTCTCTTGAAAGA
AAGTTATTTGCTCACTATCCCCAGCCTCAAGGAGCCAAGGAAGAGTCATTCACATGGAAG
GTCCGGGACTGGTCAGCCACTCTGACTTTTCTACCACATTAAATTCTCCATTACATCTCA
CTATTGGTAATGGCTTAAGTGTAAAGAGCCATGATGTGTATATTAAGCTATGTGCCACAT
ATTTATTTTTAGACTCTCCACAGCATTCATGTCAATATGGGATTAATGCCTAAACTTTGT
AAATATTGTACAGTTTGTAAATCAATGAATAAAGGTTTTGAGTGTAAAAAAAAAAAAAAA
AAAAAAA

BC007783 (IMAGE Clone ID: 4308472)

GGCACGAGGGCAAAGAGTAGTCAGTCCCTTCTTGGCTCTGCTGACACTCGAGCCCACATT
CCATCACCTGCTCCCAATCATGCAGGTCTCCACTGCTGCCCTTGCCGTCCTCCTCTGCAC
CATGGCTCTCTGCAACCAGGTCCTCTCTGCACCACTTGCTGCTGACACGCCGACCGCCTG
CTGCTTCAGCTACACCTCCCGGCAGATTCCACAGAATTTCATAGCTGACTACTTTGAGAC
GAGCAGCCAGTGCTCCAAGCCCAGTGTCATCTTCCTAACCAAGAGAGGCCGGCAGGTCTG
TGCTGACCCCAGTGAGGAGTGGGTCCAGAAATACGTCAGTGACCTGGAGCCGAGTGCCTG
AGGGGTCCAGAAGCTTCGAGGCCCAGCGACCTCAGTGGGCCCAGTGGGGAGGAGCAGGAG
CCTGAGCCTTGGGAACATGCGTGTGACCTCCACAGCTACCTCTTCTATGGACTGGTTATT

GCCAAACAGCCACACTGTGGGACTCTTCTTAACTTAAATTTTAATTTATTTATACTATTT
AGTTTTTATAATTTATTTTTGATTTCACAGTGTGTTTGTGATTGTTTGCTCTGAGAGTTC
CCCCTGTCCCCTCCACCTTCCCTCACAGTGTGTCTGGTGACAACCGAGTGGCTGTCATCG
GCCTGTGTAGGCAGTCATGGCACCAAAGCCACCAGACTGACAAATGTGTATCAGATGCTT
TTGTTCAGGGCTGTGATCGGCCTGGGGAAATAATAAAGATGTTCTTTTAAACGGTAAAAA
AAAA

X81896

AGAAAACTATTTTCTAAATATTAACACTGAAAATGTTTTGTTAGCTTTTCCTTCTTTCTC
TCCAGAAGAAACATGGATAGATGATAGCTGTTTCATTGTTTGTTTTTGTCAAGCATATTC
ACTTTCCTCCTTGTCCTCTGATTCTGAGCAAAGGGCCTCAGACTCTGAACTTCCCTCAAG
TGCCGTTGTTATGTGAACTCTTCCATTCAGATTCCAGAGAGGTTCTCATGCTCCCCCCCC
CTCCTTATTTGTAGCAATCGTAGCAACTAATTCCACTAAGTACAAGGGAGTTTTTTACAC
TCCTCCATTTTTATAGCATCTGCATTTTTTTTTTTTGTTAGGTACATGTATACACCTGCC
TGAGTATAAATACTCTCTCTACCTAATAATAACATCAACCAACATCTTTTCCAAATTAGG
GCCACAGAACAGCAACATTTGTCTGACAGTAGTATAAAGAATAATGATAGCTCTATCCTT
AAGAAGTATTTCCTTTCCTTTTTATATAGTCCCGTTAGGGTTTAAAACCATATTGATCAA
CTAGAAAGAAAAATATGAAAAGAGAAAAATATTTTAATTTAAAAATTGTAATACATTGAT
TTATAAAATGCCTTCTCTGATACTTTTGAAACAGATGTGAAAAACAGAAAAAGAAAAAAT
TGTCTGAAATGTTTATTTTGCAAAACAGTGCAATAGAATCTAGTTATGCCTTCATCACTG
TTGACAGTAAATACTGACAGCCCCTTGCAGTGTGTTAGTTTTAGATCACTCTGTTTTAGT
TGAGAGAAATGTTTTATATCATGGTTTTTATATGAATACAAATTATTTCTCAAAGATTTA
TAGCACACACTATTCTCAGGAATTCTGTATTACATGAATGCTGCTTATATATTTTCATAT
TCTAACTTGTCTTTTCAAGCAAATAACTAATATATATGTGCATGCAGTCTGCCTTGACAA
GTTGTTCCAAGCTGAAGAGCTTTCACTGTACAATGTGTGGAAAATCACCATAGATCATGG
CTGAAATAGTTTGTAATTGTCTGAGTCTGTGCACGTACTTTTAGATAAAATGCTGCTGAG
TGACTGCATGATGAGATACAACTTCTGAATGCTGCACATTCTTCCAAAATGATCCTTAGC
ACAATCTATTGTATGATGGAATGAATAGAAAACTTTTTCACTCAATAAATTATTATTTGA
TATGGTAAAAAAAAAA

BC004960 (IMAGE Clone ID: 3632495)

CCCAAGGTTGTTATATCTTCATGTCCTCATTTCTTAGGGAGGTACCTTCAGAACCAATAG
TGACCCCTAACTTCTCTGGTGGTCGGTTCCATGAAAGGCAAAGGAGTGTGAGAGAGGAGT
GGATGGTCAACCTCCCACTGCCATGGTAACATGGGTGCTGGCTGATGGGAGCAGAAAATA
ATTTAGTGAAAGTCTGTGGGGGCAGTCACAAGATGTCTGAGAAAACTGGCGAGCCAGCTG
CTGAAAACAGGGACAAGGAAGCCTCCGTGGCTGGAGCCCAAATCACACTGCAGACCCAGA
CACCGTGACCACCACCATGGACTCCAGAGAGAGCAGCTTATAGTACTCAATCAGCTGCCA
CTACCACCATCCAGAACACCAGATGTTGTAGCCATGGCTGCAGCAGGAATGGATGTCCCA
CTGTCCCTGCTCCTCGGTGTGACTTGCTCCCAAGTTCAGGGCAGGTCCATCTGATTGGCT
GAGTCTGGAATGTCTGCCTGTGCCTCAGCTGTGAGGGAGGCAGGGAAAGTAAGCCTTTTC
AGCTTCTGTCGTGGGAGGTGGGCTCTGCCTCCTACCAAGAATCAAAGGGTGGAGGATCTT
CAAACACAGGAAAAGAACCCGGATCCTGGCACCCCCAAATTTTCAGAGTCCATTTCAGAG

CATAAGAAATTGAGGGTCCAAGATCATTCATGTAAGAAGTTTAGAGGGGGAAGAAAAGAA
TGATAAACGAAAAGAACAGCAATAGTAAAGGATCTTTTCTTTGTTTCAGTAAGATGAAGA
GGCCTGAGCAGTTTCGTGGAGGGGAAGAAACAGGAAAACCTCTTCAAAAGACAAAAAGCT
GGCACTGCATTCTCTCTCTGTAGCAGGACAGAACTGTCTAAAGACAAGACCCCTTTGGCC
AAAATAAAGGAACCTGAAACATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAACCTCGGG

AK027250

AAATTTAATTAATTATAAACTCAGTCTCTTGGTTGCACCAGCCACATTTCAGATGCTCAA
TAGCCACATGTGGCTAGTGGGTACCATATTGGACAGGGCAGCTATAGAATATTTCCATCA
TTGCAGAAAGTTCTATTGGATAGTACCATAATCTTTTTATAGTAACTTGGAAATACTATT
TGATATTAGATGTTAGACCACAAAAAGAAGAAAAATGTTAGGACTATTTCAGATATAAAA
AGGAACTGAATTGTGACATAATTAGCATCTTACATTCCATACAGTTGAATACCTTATGCT
GTGACAACCATAGTTAATCATTTCAGTGCTGTTCAACATACATACCTATCAGCAGTGTGT
TTAGACCAGGGGTCTGCAAACTTTCTGTGAATGGACAAAGAGTAAATACTTTAGTAAATG
TCTTAGGCTTTGTGGCCTACATGATCTTTGTTGCAAGTACTCAACTCTGCCATTATAGAG
TTAAAGCAGCCATACACAATATATAAACAAAATGGGCATAGTTGTATTTCAGTAAAACTT
TATTTACAAAGACAGGCGGTAGGCCAGATTTGGCTTGCATGCTGTAGAGCTGTGGTCTAA
ATTTTATTCATAGACTTTCTTTGCAAATACAGTGTGAGTATTGTTCCATTTACAGTATTA
TTATTTTTTAGATACCTGGTTTTTAGATTCTTGCCTGGTAACTTTTTACTGAAAATACAA
GAATTTCGTACTGCATTTGCATCTCCGAGATTAGGGAGCACCTGTCAGGATATGTTGTTC
TATCAGGGTTACTTCTGTTGACTACCTCTTAGATTTTGATACAGTTATATTGTTGAGTTT
CATTTTCATATATTCTTGTAGTGTCTGCTTGCCTGTGACTTCTGGTAAAATAAAATAAGC
CTTTGAAAATATTTTAGCATGGTATTTAACATTTTCTAAATATTATGGCATTTTGACATA
TTTTAGTCAGCGAAGACATCTGCCCCTTTGGTGTTTCTACTTGCTTATGATTGAGATTTT
ACAAGCCCTTCAAACTCCGTTTTAAAGGAATTTATTGTAAAACATTAACTTTAATAAATT
AGTGTTTTCACAGATCAGATCATTATACTTGGAACTTCTAAATCATGCAATTTCTGAATA
AGGACATAAGGCTAGATTCATTTTTCTTAATAGAGAAAAAGGAAATTTCTGATTTATCAC
TTTTCTAGTTGATAAGTAGGATTCAAAACGTTTGATATGTAAGTATTTATATAAGACTAA
TGTAATTTAAAGTTCTGTATTATTGTGATTAATCATACAGAAATTCAGGAACTGATCAGA
AGTGAGATTCTTTTCCACATCTGGTTAATGTAGTGAGTTGACACCCTGTGGGTGGTAAAG
CATTATAAACATTTCATCTTGAACCATGATTTATACACATCTGTGTTATAAGGGAGGCTT
GAGTACATATACCAATGAAGAGATATTCAGCATTTGTCTATTTGATAAGGAATTAAATGT
CCTAGTGATTATAAAGTAAAACCACAGACCAATTTGCAAATGATCTTCAATGTTAAGCAC
TTGCTCTAAGATTAAAATTCCTTTTCTTTTTAAGGTTAAGGGTGTGTACGTATGGCAGTG
ATGTCTATGTTGAGATTAACTTATGTATTGAGGAAAATTTGAAGTTTATTTTTTCGATGA
ATAAGGCTGTCAAATGATTTAGTATAGATTAATGACATCTTTTTTAGAAATATTAAAGTG
AGTATTCCTCATTATGTCATCATTTCTGATAATTAGAGTGCTAATTTGAATGTTAGATAA
TGTTTCCACATCTATACCTATTTCTTTCTAGGGCACTTCTGACCCTGGGGCTTGGGGATG
GCCTTTAGGCCACAGTAGTGTCTGTGTTAAGTTCACTAAATGTGTATTTAATGAGAAACA
TTCCTATGTAAAAATGTGTGTATGTGAACGTATGCATACATTTTTATTGTGCACCTGTAC
ATTGTGAAGAAGTAGTTTGGAAATTTGTAAAGCACAAACCATAAAAGAGTGTGGAGTTAT
TAAATGATGTAGCACAAATGTAATGTTTAGCTTATAAAAGGTCCTTTCTATTTTCTATGG

CAAAGACTTTGACACTTGAAAAATAAAACCAATATTTGATTTATTTTTGTAAGTATTTAG
GATATTATTTTAAATAAATGATTGTCCATTATCAATAAAAAAAAAAAAAAAAAA

Sequences from Table 5 not disclosed above

NM_014298

GTCCTGAGCAGCCAACACACCAGCCCAGACAGCTGCAAGTCACCATGGACGCTGAAGGCC
TGGCGCTGCTGCTGCCGCCCGTCACCCTGGCAGCCCTGGTGGACAGCTGGCTCCGAGAGG
ACTGCCCAGGGCTCAACTACGCAGCCTTGGTCAGCGGGGCAGGCCCCTCGCAGGCGGCGC
TGTGGGCCAAATCCCCTGGGGTACTGGCAGGGCAGCCTTTCTTCGATGCCATATTTACCC
AACTCAACTGCCAAGTCTCCTGGTTCCTCCCCGAGGGATCGAAGCTGGTGCCGGTGGCCA
GAGTGGCCGAGGTCCGGGGCCCTGCCCACTGCCTGCTGCTGGGGGAACGGGTGGCCCTCA
ACACGCTGGCCCGCTGCAGTGGCATTGCCAGTGCTGCCGCCGCTGCAGTGGAGGCCGCCA
GGGGGGCCGGCTGGACTGGGCACGTGGCAGGCACGAGGAAGACCACGCCAGGCTTCCGGC
TGGTGGAGAAGTATGGGCTCCTGGTGGGCGGGGCCGCCTCGCACCGCTACGACCTGGGAG
GGCTGGTGATGTTGAAGGATAACCATGTGGTGCCCCCCGGTGGCGTGGAGAAGGCGGTGC
GGGCGGCCAGACAGGCGGCTGACTTCGCTCTGAAGGTGGAAGTGGAATGCAGCAGCCTGC
AGGAGGTCGTCCAGGCAGCTGAGGCTGGCGCCGACCTTGTCCTGCTGGACAACTTCAAGC
CAGAGGAGCTGCACCCCACGGCCACCGCGCTGAAGGCCCAGTTCCCGAGTGTGGCTGTGG
AAGCCAGTGGGGGCATCACCCTGGACAACCTCCCCCAGTTCTGCGGGCCGCACATAGACG
TCATCTCCATGGGGATGCTGACCCAGGCGGTCCCAGCCCTTGATTTCTCCCTCAAGCTGT
TTGCCAAAGAGGTGGCTCCAGTGCCCAAAATCCACTAGTCCTAAACCGGAAGAGGATGAC
ACCGGCCATGGGTTAACGTGGCTCCTCAGGACCCTCTGGGTCACACATCTTTAGGGTCAG
TGAACAATGGGGCACATTTGGCACTAGCTTGAGCCCAACTCTGGCTCTGCCACCTGCTGC
TCCTGTGACCTGTCAGGGCTGACTTCACCTCTGCTCATCTCAGTTTCCTAATCTGTAAAA
TGGGTCTAATAAAGGATCAACCAAAAAAAAAAAAAAAAAAA

AF033199

CGGGGCATGCTGCTTCCCTTCACCTTCCACCATGATTGTAAGTTTCCTGAGGCCTCCCCA
GGTGTGCTTCTGTACAGCCTGTGGAATGTTACCAAAGACGTTGGAAGAGGTGGCTATGGG
ACATCACCTGGGAGAAGTGGAAGCAAATGGACACTGTTCAGAAGTCCATATACAGAAACA
TACTTGGAAAAATATAGAAACCTGGTTTTGCTAGATGGGAAGCTTGCAGCTGGGGCCAAG
ACATCAAGAGTAGAGCAGCAGGACATTTCAAAAGAAGATTAACTCAAAGATTAGAGATGG
AAGAACTTGCAAAGAGAAAGTCTGTACCGGAAGAAATCTGGAAATCTAGAGGCCAGTTTA
AGAATCAGCAGCTAAACAAGGAGAATAATCTAGGGCAAGAGATAGCTACCTGCACAAAAA
TTCCTACCAGAAAAAGAGACATAGAATCTAATGAATTTGTGAAAAATTTTACTGTAAGAT
CAATACTTGTTGCAGAACAGATAGATCCTATGGAAGAGAATTGTCATAAATATGGTACAT
GTTGAAAGATGCTCAAACAAAACTCAGATTTAATTATACAAAGAAAGTATGATGGAAAAA
AAAAAACCTTGTAAATATAGTGAATGTGGGAGAACCTTCAGAGGCCACATCACTCTTGTT
CAGCATCAAATAACTCATTGTGGAGAGAGACCCTGTAAATGTACTGAGTGTAGAAAGGGA
TTTAATCAGAGTTCCCACTTAAGAAATAATCAGAGAAAAACTCTTTCAGGAGAAAAGCCC

TACAAATGCAGTGAGTGTGGGAAGGCCTTCAGTTATTGCTTAGTTCTTAATCAACACCAG
AGAATTCACAGTGGAGAGAAACCTTATGAGGGTACTGAATGTGGCAAGACATTCATTCAG
TCGTACATACCTTACTCAGCATCAAAGAATTCACACACTGGTGAGAAGCCCTATACATGT
CTTGAATGTGGAAGGCTTTTTAGTCAGAACACACATCTTACTCTACATCAGAGAATCCAT
ACTGGAGAGAAACCTTATGAATGCAATGAATGTGGTAGGTCCTTTAGTCAGACTGCACAT
CTTACTCAACATCAAAGAATGTATACAGGAGAAAAACTCTATGAATGTAATGAATGTGAG
AAAGCCTTCCATGATCACTCAGCTCTTATTCAACATCATATTGTCCATACTGCAGAGAAA
CCCTATGATATCATGACTGGGAAAACTTTCAGTTACTGTTCAGACCTCATTCAACATCAG
AGAATGCACACTGGAGAGAAACCATACAAATGCAATGAATGTGGGAATGCCTTTAGTGAT
TGTTCATCCCTTATTCAGCATCAAAGAACTCACACTGGAGAAGAGCCTTATGAATGTAAG
CAATGTGGAAAAGCCTTTAGCAGAAGCACATACCTTACTCAACATCAGAGAAGTCACGCA
GGAGAGAAACAGTATAAATGCAATGAATGTGAGAAAACTTTCAGCCTGAGTTCATTCCTT
ACACAGCATATGAGGGTTCAGACTGGAGAAAAACCCTACAAATATAATGAATATGGAAAA
GCTTTTAGTGACTGCTCAGGACATTTTCAGAGAACTCACACTGGAGAGAAGCCCTGTGAA
TGTAATGACTGTGGGAAACCTTTCAGTTTCTGTTCAGCCCTAATTCAACATAAGAGAATT
CATACCAGAAAGAAGCCCTGACTGTACCTTCATACCAGTAAATGCACTGACTGTGGAAAA
GCCTTCAGTGATTGGTTAGCACTTGTTCAACATCAGATAACTCAACACTGGAGAAAAACC
GTATAAATGTACTGAATGTGGAAAAGCCTTCAGTTGGAGTACAGACCTCAAAAATCACCA
GAAAACTCATACTAGTGAAAAATCCTATAAATGTAATGAATGTAGAAAGGCCTTTAGTTA
CTGCTCTGGTCTTATTCAATGTCAGGTCATTCATACTATAGAAAAACCTTATGAATACGG
TAAATGTGGCAAAGCCTTTAGGCAGAGGACAGACCTTAAAAAACATCAGAAAATGCATAC
CGAAGAGAAACCCTATGAATGTAATGAATGTGGGAAAGCCTTTAGCCAGAGCACATATCT
TACAAAACACCAAAAAATTCATAGTGAAGAGAAATCAAATATACATACTGAGTGTGGGGA
AACCATTAGACAAAACTCTTCTTTTTACAACAATAAAACCTCACACTGGAGAGTTCTCTG
AATGCCTTAAGAATTTGGTTAATATGGAGACCCTTCCCAGGGAAACAGAAGGAGGATCGT
GAAAACCGTTGACTACTTGAATGATCACATGGTTTAGTGGAGAGAGCATGATTCTGGGTT
TTAAAAGTCATGGATCTCAATCTCAGCTCCTATTACTAACTAGATCTTTTACTTTGGGGT
AAGTCACTTCATATCTTTAGGCCTTAATTTCCTCATCTGAAAACTGGAAGGCCTGACTTG
ACTTGTTGAGCTTAAGATCCTCAATTATTATATTTACTAGGAATTCAAGTTTCTATAGAT
GTGGTTCAGAATTGTGACTTATTTATTGTACATCAGGTGTGATTCACAAGTGAGCTTGTA
GTAGTTATTAAGGAGTCAATAAAGATATGATATAAAAAAAAAAAAAAAAA

AI688494 (IMAGE Clone ID: 2330499)

CATTTCATCTTCATTGGATAGTGTTACATAGTAATATATTTATGTTTTCTTTTAATCATT
TCATAACTTGGAAAATACTAACATAGTCAAAACTCTAGGGTAGGTGATACATGAGTTTCT
GTAGTAATCTGGTTGGAGACATGTTGTAATTCTGTATATATATGTACATTTATCCCATGC
ATGTTATGCCTAAACTAAGACGGATACCCCTGAATTAAGAGGTGCTGTTATACATTGACC
AGGCTTAAGAATATCTCTTTAAAGTGTGTCGACATTTAATTGACCTTTGGAAGTTCATTC
TGTTAATCATACTCAAAGTGCTAAAGCTATGGTTGACTGCTCTGGTGTTTTTATATTCAT
TCGTGCTTTAGCATATAAATTCTTCAGCATAATTGCTACTTATTTAGCAAGAGTTTCCTT
TATTTGAAAATGTGAGTTGTGCTTGTATTTTTGTGTCTTTCTTTCTTTCTTTCTTTTTTT
AAACTTTGCTTCAGGCTGGGTAGTGGTAGAGGTTTGAATTAAAATGTTTTCCTGTCAGTA
AAAAAAAAAAA

AL157459

GAGCGAGCCCAGCAGCTTGCCCTTGACAGGTGGGGGCTGGCTGGGGCCTTAATGTGAAAA
GACAGTGGCAGGCAGCTGGAGTAGAGCGAGCCCAGCAGCCCTAAAAGGCTGCCTTCATGG
CCATCTAGCCCCAGTTCAGGGCAGCATCCATAGCCCACAAGCCAGCGTGGGTGGGGCGGG
GGTGGTCCCACAGCTGGGTTCCACCTGAAGAGCCTCCGTGCCTCGGAGCAGGAGAGGCAG
GCTATGGCTGTCACCCTCCCTCCTGCCTGTGTCCCAGTGAGAACTGACCTGAGTCCCCTT
CCAAACCCAGACCCACCTCCTGCCCCAGGCCCACTGAAGCATGTTCCATTTCTAAAAAGC
CCAGAGTTCAGTGTGTCCCAAGGAAAACCCAAAGTGGAGGTGCTCAGGTCCAGGGGAGTC
CAGTGGGCAGGACCCTTGGCAGGCAAGCCCCTCCCTTCACTCCCAGGACCTACCTTCTGC
TAGTAAAGGACTGGCTTCATTCTAATTATGGCCCACAGACTGCCCCGGAGACCTGGAGGA
CAGCAGTGCTGGCACTTGGGTGTCCATGGGCCCGTCTGCCGGCTCTGCCTGTGCTGCAAG
TGTTGGCCGTGGGTCCAGCCAACAACTCCCTACGTCCTGTGTGGGGCCCTGCCCAAGTGG
ATGAGGCATTCCTTGAGGAGTATCATTTTCCCTGACAATCCCCATCACCTTTAGGGGTTC
CCTGCTTGGCTCCTTTCCAGCTGAAAAACTAGACCTGTGCCATTGGGGAAGCTGGACAAA
GTCTAGGGGGCCCGCCTGGTAGAGGGTCCCGGGAAGCTGGATCTGTCAGCCTCGGCCCTG
AGGCCCCTGTTAACTCAAGACTGTGAGCTGCCTCTAGGTGGTCACGTCTGGGAGCTAGCT
TGTATGGCTTCTGACCAGTATCAGGATTTCTGTTCTGAGAGCAGCGTGGGCAGCAAGGCA
GGGCAGCCCAGAGGTGGCAGCGGCAGGCAATCTGGTCACTAGGTCTTTGTGATGCCAAAA
ATAAAAGAGGGTGGGGTGGGTGCTTTCTGTTCCTCTGATTGGATGGAGTCCGCCAGCAGG
CATGGGGCTACATTCCAGTGCCTGACTATAGGGAGGCACTCCTGATTCCATGGAGCAGCC
CGGACTTTGAGAATGGGCTCTGGTTTGCGGGGGGCAGGCGTACCAGACTGCAAGACCCCC
CAGTACCTCACCGTGCCAAATAGGAAGAGGTGGCCTTGGTGTAGCCAAATGGATCTTTTT
AACAGTGTGCCTTTGGGGAGGGACCCATGTCCATGGCTTCGTTGAGGGCCATCCATATGC
CAGCTGGGGGCCAGCCCACAGTGGCCATATTGGCTGCAGCAGGAATGGTGCCCACCTCGG
CGAATTGAAGGGCTAAGAGTCCCAGATAGCTAGGCCAGAGCTGGAAGCAGACAGTAAGGG
GAAGAGCTGCTCCCACAGGAGAGGGAGAGATTCCAGCTCACTGCGCAGCCTGGGAGGAGG
CGTGGATCCTGGCACGCTGAGCCTCAGGCACCAGCCTCCCTGTGCTCGACAGCAAAGTCT
TGACTCCTTCCTGCTGAGCACTGTGCTACCTTCACTGCTCCAAAGCCAGACTAACAGCTC
TCCAAGCCCTTGGGGTGACTCGGCTTCCAGGAGCTGTTGGAGAAATGAGGATGTCTGTCC
CTGTCTGCCTGGGCAGGCCAGATTCCTCCCCAGCAGCCGGGTCTCTCCAGACCCTGATTC
GGTGCCTTTCTGTTTACCAGCTACTTCAATCCCAAAGTTTGAATCTGCAGATACCTTACT
CCCAGCCACTTTGCCTTCTTACTGTGTTGTGTGTTTTTCCTGGTGCTTCAAGAGCGTGTG
CAGGGCAAGTGCCGTCACTGGGAACTGCACCAGATGCTCAGACTTGGTTGTCTTATGTTT
ACCAATAAATAAAAGTAGACTTTTTCTATTTTTATTTGCTGCTATTTGTGTGTGTGTTTG
TGTTTGTGTAGCTAGGTATCTGGCACTTCTGACGATGCATTGTTGCTTTTTTCCCGAAGG
TCCCGCAGGAACTGTGGCAATGGTGTGTGTGTGAAATGGTGTGTTAACCGCGTTTTGTTT
GCTCCTGTATTGAATAGGAAGCAGTGGCCAGTCTGTCTTCCTTAGAGATGTTAGCATATT
TTTATATGTATATATTTTGTACCAAAAAAGAGTGTTCCTTGTTTTGGTTACACTCGAAAT
TCTGACCTAGCTGGAGAGGGCTCTGGGCCGAGAGCTTTCACTAAGGGGAGACTTCAGGGG
AGGATCAAGCTTTGAACCAAAGCCAATCACTGGCTTGATTTGTGTTTTTTAATTAAAAAA
AAAATCATTCATGTATGCCACTTCTAAAAAAAAAAAAAAAAAAAAAAAA

BC002480 (IMAGE Clone ID: 3350037)

GGCACGAGGCTGAGACCGGTGCGCCGCGCGCTAGTGGCCGCTCTTCCGCGGGCTAGCGGG
CGGTGGGGGCGCCAGCAGCGCGGAAGGCGGGCACGCGGGCCATGGCTCCCTGGGCGGAGG
CCGAGCACTCGGCGCTGAACCCGCTGCGCGCGGTGTGGCTCACGCTGACCGCCGCCTTCC
TGCTGACCCTACTGCTGCAGCTCCTGCCGCCCGGCCTGCTCCCGGGCTGCGCGATCTTCC
AGGACCTGATCCGCTATGGGAAAACCAAGTGTGGGGAGCCGTCGCGCCCCGCCGCCTGCC
GAGCCTTTGATGTCCCCAAGAGATATTTTTCCCACTTTTATATCATCTCAGTGCTGTGGA
ATGGCTTCCTGCTTTGGTGCCTTACTCAATCTCTGTTCCTGGGAGCACCTTTTCCAAGCT
GGCTTCATGGTTTGCTCAGAATTCTCGGGGCGGCACAGTTCCAGGGAGGGGAGCTGGCAC
TGTCTGCATTCTTAGTGCTAGTATTTCTGTGGCTGCACAGCTTACGAAGACTCTTCGAGT
GCCTCTACGTCAGTGTCTTCTCCAATGTCATGATTCACGTCGTGCAGTACTGTTTTGGAC
TTGTCTATTATGTCCTTGTTGGCCTAACTGTGCTGAGCCAAGTGCCAATGGATGGCAGGA
ATGCCTACATAACAGGGAAAAATCTATTGATGCAAGCACGGTGGTTCCATATTCTTGGGA
TGATGATGTTCATCTGGTCATCTGCCCATCAGTATAAGTGCCATGTTATTCTCGGCAATC
TCAGGAAAAATAAAGCAGGAGTGGTCATTCACTGTAACCACAGGATCCCATTTGGAGACT
GGTTTGAATATGTTTCTTCCCCTAACTACTTAGCAGAGCTGATGATCTACGTTTCCATGG
CCGTCACCTTTGGGTTCCACAACTTAACTTGGTGGCTAGTGGTGACAAATGTCTTCTTTA
ATCAGGCCCTGTCTGCCTTTCTCAGCCACCAATTCTACAAAAGCAAATTTGTCTCTTACC
CGAAGCATAGGAAAGCTTTCCTACCATTTTTGTTTTAAGTTAACCTCAGTCATGAAGAAT
GCAAACCAGGTGATGGTTTCAATGCCTAAGGACAGTGAAGTCTGGAGCCCAAAGTACAGT
TTCAGCAAAGCTGTTTGAAACTCTCCATTCCATTTCTATACCCCACAAGTTTTCACTGAA
TGAGCATGGCAGTGCCACTCAAGAAAATGAATCTCCAAAGTATCTTCAAAGAATAAATAC
TAATGGCAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA